

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 13, 2003, 04:37:24 ; Search time 1755.93 Seconds  
(without alignments)  
7460.491 Million cell updates/sec

Title: US-09-852-261-3  
Perfect score: 539  
Sequence: 1 ggaccagagaccctttgcgg.....agtaaacattcccgaattc 539

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_htc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_htc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*

28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result			%				ID	Description
	No.	Score	Match	Length	DB	Query		
c	1	501.6	93.1	549	9	AI169253	AI169253	EST215088
	2	480.4	89.1	614	14	CD373004	CD373004	UI-R-GR0-
c	3	469.2	87.1	558	9	AI503976	AI503976	vm43d08.x
c	4	453.4	84.1	623	9	AW146128	AW146128	um37e10.x
c	5	450.4	83.6	468	9	AI169770	AI169770	EST215669
c	6	446.8	82.9	558	9	AI265629	AI265629	uj04b07.x
c	7	437	81.1	653	13	BQ200567	BQ200567	UI-R-DZ1-
	8	431.8	80.1	594	10	BF383724	BF383724	602044632
	9	425.2	78.9	816	9	AI119218	AI119218	ue94h02.y
c	10	423	78.5	499	9	AW495481	AW495481	UI-M-BH3-
c	11	421.2	78.1	642	9	AI876493	AI876493	uj59b10.x
c	12	402.6	74.7	525	9	AA963258	AA963258	UI-R-E1-g
c	13	401.6	74.5	500	9	AA945553	AA945553	EST201052
c	14	399.4	74.1	525	9	AI599751	AI599751	EST251454
c	15	395	73.3	499	12	BI294072	BI294072	UI-R-DK0-
c	16	394.2	73.1	502	9	AI104669	AI104669	EST213958
c	17	367	68.1	470	9	AI233293	AI233293	EST229981
c	18	362.6	67.3	521	9	AW493459	AW493459	UI-M-BH3-
	19	362.6	67.3	621	12	BI221656	BI221656	602936980
	20	362.6	67.3	1658	11	AK081019	AK081019	Mus muscu
	21	361.6	67.1	559	12	BI715603	BI715603	ic34h10.y
c	22	361.6	67.1	595	9	AI573421	AI573421	mo04b11.x
	23	361.6	67.1	602	13	BU590710	BU590710	AGENCOURT
c	24	360	66.8	499	12	BI676839	BI676839	ic56a08.x
	25	355.2	65.9	559	12	BI715465	BI715465	ic33b09.y
c	26	352.4	65.4	477	9	AA800127	AA800127	EST189624
c	27	350.4	65.0	561	12	BI714874	BI714874	ic33b09.x
c	28	349.2	64.8	470	9	AA945027	AA945027	EST200526
c	29	348	64.6	564	12	BI714981	BI714981	ic34h10.x
	30	346.4	64.3	558	12	BI715475	BI715475	ic33c08.y
c	31	341.4	63.3	464	10	BE104013	BE104013	UI-R-BX0-
c	32	339.4	63.0	448	9	AW252120	AW252120	UI-R-BJ0-
	33	338	62.7	882	9	AI604642	AI604642	vm43d08.y
c	34	335.6	62.3	464	12	BI277678	BI277678	UI-R-CZ0-
c	35	333.6	61.9	452	9	AW525416	AW525416	UI-R-BJ0p
	36	329.6	61.2	665	9	AA690767	AA690767	vu57d12.r
	37	329.4	61.1	799	9	AI314558	AI314558	uj48d07.y
c	38	327.4	60.7	460	9	AA924219	AA924219	UI-R-A1-d
c	39	326	60.5	429	9	AI101163	AI101163	EST210452
	40	323.2	60.0	773	12	BI144500	BI144500	602908689
c	41	321.4	59.6	637	9	AW413016	AW413016	uq49h08.x
	42	316	58.6	949	14	CB589117	CB589117	AGENCOURT
c	43	311.6	57.8	445	10	BG071465	BG071465	H3099C09-
	44	310.4	57.6	473	9	AA451360	AA451360	vf84g03.r
c	45	300.4	55.7	474	9	AI526955	AI526955	uj48d07.x

## ALIGNMENTS

## RESULT 1

AI169253/c

LOCUS AI169253 549 bp mRNA linear EST 08-JAN-1999  
 DEFINITION EST215088 Normalized rat kidney, Bento Soares Rattus sp. cDNA clone  
 RKIBP33 3' end, mRNA sequence.

ACCESSION AI169253

VERSION AI169253.1 GI:4134375

KEYWORDS EST.

SOURCE Rattus sp.

ORGANISM Rattus sp.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.

REFERENCE 1 (bases 1 to 549)

AUTHORS Lee, N.H., Glodek, A., Chandra, I., Mason, T.M., Quackenbush, J.,  
 Kerlavage, A.R. and Adams, M.D.

TITLE Rat Genome Project: Generation of a Rat EST (REST) Catalog & Rat  
 Gene Index

JOURNAL Unpublished

COMMENT On Oct 6, 1998 this sequence version replaced gi:3705561.

Other\_ESTs: TC50779.

Contact: Lee, NH

The Institute for Genomic Research

9712, Medical Center Drive, Rockville, MD 20850, USA

Tel: (301)-838-3529

Fax: (301)-838-0208

Email: nhlee@tigr.org

Seq primer: M13-21.

FEATURES Location/Qualifiers

source

1..549

/organism="Rattus sp."

/mol\_type="mRNA"

/db\_xref="taxon:10118"

/clone="RKIBP33"

/clone\_lib="Normalized rat kidney, Bento Soares"

/note="Organ: kidney; Vector: pT7T3Pac; Site\_1: EcoRI;

Site\_2: NotI"

BASE COUNT 112 a 140 c 133 g 164 t

ORIGIN

Query Match 93.1%; Score 501.6; DB 9; Length 549;

Best Local Similarity 97.3%; Pred. No. 3.7e-107;

Matches 510; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

```

Qy      8 AGACCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTCGTGTGTGGACCAAGGGGCT 67
          |||
Db      549 AGACCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAATTCGTGTGTGGACCAAGGGGCT 490

Qy      68 TTTACTTCAACAAGCCACAGTCTATGGCTCCAGCATTCGGAGGGCACCACAGACGGGCA 127
          |||
Db      489 TTTACTTCAACAAGCCACAGGCTATGGCTCCAGCATTCGAAGGGCACCACAGACGGGCA 430

```

Qy 128 TTGTGGATGAGTGTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTACTGTGTCC 187  
 |||||  
 Db 429 TTGTGGATGAGTGTAGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTACTGTGCTC 370  
 Qy 188 GCTGCAAGCCTACAAAGTCAGCTCGTTCCATCCGGGCCCAGCGCCACACTGACATGCCCCA 247  
 |||||  
 Db 369 CGCTGAAGCCTACAAAGTCAGCTCGTTCCATCCGGGCCCAGCGCCACATTGACATGCCCCA 310  
 Qy 248 AGACTCAGAAGTCCCAGCCCCTATCGACACACAAGAAAAGGAAGCTGCAAAGGAGAAGGA 307  
 |||||  
 Db 309 AGACTCAGAAGTCCCAGCCCCTATCGACACACAAGAAAAGGAAGCTGCAAAGGAGAAGGA 250  
 Qy 308 AAGGAAGTACACTTGAAGAACACAAGTAGAGGAAGTGCAGGAAACAAGACCTACAGAATG 367  
 |||||  
 Db 249 AAGGAAGTACACTTGAAGAACACAAGTAGAGGAAGTGCAGGAAACAAGACTTACAGAATG 190  
 Qy 368 TAGGAGGAGCCTCCCGAGGAACAGAAAATGCCACGTACCCGCAAGATCCTTTGCTGCTTG 427  
 |||||  
 Db 189 TAGGAGGAGCCTCCCGAGGAACAGAAAATGCCACGTACCCGCAAGATCCTTTGCTGCTTG 130  
 Qy 428 AGCAACCTGCAAAACATCGGAACACCTGCCAAATATCAATAATGAGTTCAATATCATTTTC 487  
 |||||  
 Db 129 AGCAACCTGCAAAACATCGGAACACCTGCCAAATATCAATAATGAGTTCAATACCATTTC 70  
 Qy 488 AGAGATGGGCATTTCCCTCAATGAAATACACAAGTAAACATTCC 531  
 |||||  
 Db 69 AGAGATGGGCATTTCCCTCAATGAAATACACAAGTAAACATTCC 26

## RESULT 2

CD373004

LOCUS CD373004 614 bp mRNA linear EST 29-MAY-2003

DEFINITION UI-R-GR0-csv-j-17-0-UI.r1 UI-R-GR0 Rattus norvegicus cDNA clone  
 UI-R-GR0-csv-j-17-0-UI 5', mRNA sequence.

ACCESSION CD373004

VERSION CD373004.1 GI:31157094

KEYWORDS EST.

SOURCE Rattus norvegicus (Norway rat)

ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.

REFERENCE 1 (bases 1 to 614)

AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.

TITLE Normalization and subtraction: two approaches to facilitate gene  
 discovery

JOURNAL Genome Res. 6 (9), 791-806 (1996)

MEDLINE 97044477

PUBMED 8889548

COMMENT Contact: Soares, MB

Coordinated Laboratory for Computational Genomics

University of Iowa

375 Newton Road , 4156 MEBRF, Iowa City, IA 52242, USA

Tel: 319 335 8250

Fax: 319 335 9565

Email: bento-soares@uiowa.edu

Tissue Procurement: James Lin, University of Iowa





Db 356 ATGCCCAAGACTCAGAAGTCCCAGCCCCTATCGACACACAAGAAAAGGAAGCTGCAAAGG 415

Qy 301 AGAAGGAAAGGAAGTACACTTGAAGAACACAAGTAGAGGAAGTGCAGGAAACAAGACCTA 360  
 |||

Db 416 AGAAGGAAAGGAAGTACACTTGAAGAACACAAGTAGAGGAAGTGCAGGAAACAAGACCTA 475

Qy 361 CAGAATGTAGGAGGAGCCTCCCGAGGAACAGAAAATGCCACGTCACCGCAAGATCCTTTG 420  
 |||

Db 476 CAGAATGTAGGAGGAGCCTCCCGAGGAACAGAAAATTCCACGTCACCGCATGATCCTTTG 535

Qy 421 CTGCTTGAGCAACCTGCAAAACATCGGAACACCTGCCAAATATCAATAATGAGTTCAATA 480  
 |||

Db 536 CTGCTTGAGCAACCTGCANAACATCGGAACACCTGCCAAATATCAATAATGAGTTCAATA 595

Qy 481 TCATTTTCAGAGATGGGCAT 499  
 |||

Db 596 CCATTTTCAGAGATGGGCAT 614

# RESULT 3

AI503976/c

LOCUS AI503976 558 bp mRNA linear EST 11-MAR-1999

DEFINITION vm43d08.x1 Stratagene mouse diaphragm (#937303) Mus musculus cDNA clone IMAGE:1001007 3' similar to gb:X04482 Mouse mRNA for preproinsulin-like growth factor IB (MOUSE);, mRNA sequence.

ACCESSION AI503976

VERSION AI503976.1 GI:4401827

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 558)

AUTHORS Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.

TITLE The WashU-NCI Mouse EST Project 1999

JOURNAL Unpublished

COMMENT Contact: Marra M/WashU-NCI Mouse EST Project 1999  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810

Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.  
 MGI:565223

This clone was previously sequenced on the 5' end only, this new data is from the 3' end

High quality sequence stop: 440.

FEATURES Location/Qualifiers

source 1..558  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10090"

```

/clone="IMAGE:1001007"
/tissue_type="diaphragm"
/dev_stage="adult"
/lab_host="SOLR (kanamycin resistant)"
/clone_lib="Stratagene mouse diaphragm (#937303)"
/note="Organ: diaphragm; Vector: pBluescript SK-; Site_1:
EcoRI; Site_2: XhoI; Cloned unidirectionally from mRNA
prepared from diaphragm muscle. Primer: Oligo dT. Average
insert size: 1.5 kb. Uni-ZAP XR Vector; ~5' adaptor
sequence: 5' GAATTCGGCACGAG 3' ~3' adaptor sequence: 5'
CTCGAGTTTTTTTTTTTTTTTTTTT 3'"

```

BASE COUNT        103 a        133 c        149 g        173 t  
ORIGIN

Query Match                    87.1%;    Score 469.2;    DB 9;    Length 558;  
Best Local Similarity    92.8%;    Pred. No. 1.5e-99;  
Matches 492;    Conservative        0;    Mismatches    38;    Indels        0;    Gaps        0;

```

Qy      1 GGACCAGAGACCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTTCGTGTGTGGACCA 60
      |||
Db    530 GGACCAGAGACCCTTTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTTCGTGTGTGGACCG 471

Qy      61 AGGGGCTTTTACTTCAACAAGCCACAGTCTATGGCTCCAGCATTTCGGAGGGCACCACAG 120
      |||
Db    470 AGGGGCTTTTACTTCAACAAGCCACAGGCTATGGCTCCAGCATTTCGGAGGGCACCTCAG 411

Qy     121 ACGGGCATTGTGGATGAGTGTGTGCTTCCGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 180
      ||
Db    410 ACAGGCATTGTGGATGAGTGTGTGCTTCCGAGCTGTGATCTGAGGAGACTGGAGATGTAC 351

Qy     181 TGTGTCCGCTGCAAGCCTACAAAGTCAGCTCGTTCCATCCGGGCCCAGCGCCCACTGAC 240
      ||||
Db    350 TGTGCCCCACTGAAGCCTACAAAAGCAGCCCGCTCTATCCGTGCCCAGCGCCCACTGAC 291

Qy     241 ATGCCCCAAGACTCAGAAGTCCCAGCCCCTATCGACACACAAGAAAAGGAAGCTGCAAAGG 300
      |||
Db    290 ATGCCCCAAGACTCAGAAGTCCCCGTCCCTATCGACAAAACAAGAAAACGAAGCTGCAAAGG 231

Qy     301 AGAAGGAAAGGAAGTACACTTGAAGAACACAAGTAGAGGAAGTGCAGGAAACAAGACCTA 360
      |||
Db    230 AGAAGGAAAGGAAGTACATTTGAAGAACACAAGTAGAGGAAGTGCAGGAAACAAGACCTA 171

Qy     361 CAGAATGTAGGAGGAGCCTCCCGAGGAACAGAAAATGCCACGTCACCGCAAGATCCTTTG 420
      |||
Db    170 CAGAATGTAGGAGGAGCCTCCACGGAGCAGAAAATGCCACATCACCGCAGGATCCTTTG 111

Qy     421 CTGCTTGAGCAACCTGCAAAACATCGGAACACCTGCCAAATATCAATAATGAGTTCAATA 480
      |||
Db    110 CTGCTTGAGCAACCTGCAAAACATCGAAACACCTACCAAATAACAATAATAAGTCCAATA 51

Qy     481 TCATTTTCAGAGATGGGCATTTCCCTCAATGAAATACACAAGTAAACATTC 530
      |||
Db     50 ACATTACAAAGATGGGCATTTCCCCCAATGAAATATACAAGTAAACATTC 1

```

LOCUS AW146128 623 bp mRNA linear EST 10-OCT-2000  
 DEFINITION um37e10.x1 Sugano mouse embryo mewa Mus musculus cDNA clone  
 IMAGE:2247498 3' similar to gb:X04482 Mouse mRNA for  
 preproinsulin-like growth factor IB (MOUSE);, mRNA sequence.  
 ACCESSION AW146128  
 VERSION AW146128.1 GI:6167864  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 623)  
 AUTHORS Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,  
 Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person  
 ,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter  
 ,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,  
 Waterston,R. and Wilson,R.  
 TITLE The WashU-NCI Mouse EST Project 1999  
 JOURNAL Unpublished  
 COMMENT Contact: Marra M/WashU-NCI Mouse EST Project 1999  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: mouseest@watson.wustl.edu  
 This clone is available royalty-free through LLNL ; contact the  
 IMAGE Consortium (info@image.llnl.gov) for further information.  
 MGI:1006958  
 Seq primer: custom primer used  
 High quality sequence stop: 499.  
 FEATURES Location/Qualifiers  
 source 1. .623  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:2247498"  
 /dev\_stage="embryo, 14 dpc"  
 /lab\_host="DH10B"  
 /clone\_lib="Sugano mouse embryo mewa"  
 /note="Vector: pME18S-FL3; Site\_1: DraIII (CACTGTGTG);  
 Site\_2: DraIII (CACCATGTG); 1st strand cDNA was primed  
 with an oligo(dT) primer [ATGTGGCCTTTTTTTTTTTTTTTTTT];  
 double-stranded cDNA was ligated to a DraIII adaptor  
 [TGTGGCCTACTGG], digested and cloned into distinct DraIII  
 sites of the pME18S-FL3 vector (5' site CACTGTGTG, 3' site  
 CACCATGTG). XhoI should be used to isolate the cDNA  
 insert. Size selection was performed to exclude fragments  
 <1.5kb. Library constructed by Dr. Sumio Sugano  
 (University of Tokyo Institute of Medical Science).  
 Custom primers for sequencing: 5' end primer  
 CTTCTGCTCTAAAAGCTGCG and 3' end primer  
 CGACCTGCAGCTCGAGCACA."  
 BASE COUNT 123 a 138 c 170 g 191 t 1 others  
 ORIGIN

Query Match

84.1%; Score 453.4; DB 9; Length 623;

Best Local Similarity 92.1%; Pred. No. 7.5e-96;  
Matches 489; Conservative 0; Mismatches 41; Indels 1; Gaps 1;

```

Qy      1 GGACCAGAGACCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTCGTGTGTGGACCA 60
      |||
Db      541 GGACCAGAGACCCTTTTCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGACCG 482

Qy      61 AGGGGCTTTTACTTCAACAAGCCCACAGTCTATGGCTCCAGCATTTCGGAGGGCACCACAG 120
      |||
Db      481 AGGGGCTTTTACTTCAACAAGCCCACAGGCTATGGCTCCAGCATTTCGGAGGGCACCTCAG 422

Qy      121 ACGGGCATTGTGGATGAGTGTGTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 180
      ||
Db      421 ACAGGCATTGTGGATGAGTGTGTGCTTCCGGAGCTGTGATCTGAGGAGACTGGAAATGTAC 362

Qy      181 TGTGTCCGCTGCAAGCCTACAAAGTCAGCTCGTTCCATCCGGGGCCAGCGCCACACTGAC 240
      ||||
Db      361 TGTGCCCACTGAAGCCTACAAAGCAGCCCGCTCTATCCGTGCCAGCGCCACACTGAC 302

Qy      241 ATGCCCAAGACTCAGAAGTCCCAGCCCCTATCGACACACAAGAAAAGGAAGCTGCAAAGG 300
      |||
Db      301 ATGCCCAAGACTCAGAAGTCCCCGTCCCTATCGACAAACAAGAAAACGAAGCTGCAAAGG 242

Qy      301 AGAAGGAAAGGAAGTACACTTGAAGAACACAAGTAGAGGAAGTGCAGGAAACAAGACCTA 360
      |||
Db      241 AGAAGGAAAGGAAGTACATTGAAGAACCCAAGTAGAGGAAGTGCAGGAAACAAGACCTA 182

Qy      361 CAGAATGTAGGAGGAGCCTCCCGAGGAACAGAAAATGCCACGTCACCGCAAGATCCTTTG 420
      |||
Db      181 CAGAATGTAGGAGGAGCCTCCACGGAGCAGAAAATGCCACATCACCGCAGGATCCTTTG 122

Qy      421 CTGCTTGAGCAACCTGCAAAACATCGGAACACCTGCCAAATATCAATAATGAGTTCAATA 480
      |||
Db      121 CTGCTTGAGCAACCTGCAAAACATCGAAACCCCTACCAAATAACAATAATAAGTCCAATA 62

Qy      481 TCATTTTCAGAGATGGGCATTTCCCTCAATGAAATACACAAGTAAACATTCC 531
      ||||
Db      61 ACATTACAAAGATGGGCATTTCCC-CAATGAAATATACAAGTAAACATTCC 12

```

# RESULT 5

AI169770/c

LOCUS AI169770 468 bp mRNA linear EST 20-JAN-1999

DEFINITION EST215669 Normalized rat liver, Bento Soares Rattus sp. cDNA clone  
RLIAT07 3' end, mRNA sequence.

ACCESSION AI169770

VERSION AI169770.1 GI:3709810

KEYWORDS EST.

SOURCE Rattus sp.

ORGANISM Rattus sp.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

REFERENCE 1 (bases 1 to 468)

AUTHORS Lee,N.H., Glodek,A., Chandra,I., Mason,T.M., Quackenbush,J.,  
Kerlavage,A.R. and Adams,M.D.

TITLE Rat Genome Project: Generation of a Rat EST (REST) Catalog & Rat

Gene Index  
 JOURNAL Unpublished  
 COMMENT Other\_ESTs: TC50779  
 Contact: Lee, NH  
 The Institute for Genomic Research  
 9712, Medical Center Drive, Rockville, MD 20850, USA  
 Tel: (301)-838-3529  
 Fax: (301)-838-0208  
 Email: nhlee@tigr.org  
 Seq primer: M13-21.

FEATURES Location/Qualifiers  
 source 1. .468  
 /organism="Rattus sp."  
 /mol\_type="mRNA"  
 /db\_xref="ATCC (inhost):2027570"  
 /db\_xref="taxon:10118"  
 /clone="RLIAT07"  
 /clone\_lib="Normalized rat liver, Bento Soares"  
 /note="Organ: liver; Vector: pT7T3Pac; Site\_1: EcoRI;  
 Site\_2: NotI"

BASE COUNT 85 a 115 c 119 g 149 t  
 ORIGIN

Query Match 83.6%; Score 450.4; DB 9; Length 468;  
 Best Local Similarity 97.6%; Pred. No. 3.6e-95;  
 Matches 457; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

```

Qy      63 GGGCTTTTACTTCAACAAGCCACAGTCTATGGCTCCAGCATTTCGGAGGGCACCACAGAC 122
      |||
Db      468 GGGCTTTTACTTCAACAAGCCACAGGCTATGGCTCCAGCATTTCGGAGGGCACCACAGAC 409

Qy     123 GGGCATTGTGGATGAGTGTGTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTACTG 182
      |||
Db     408 GGGCATTGTGGATGAGTGTGTGCTCCCGAGCTGTGATCTGAGGAGGTTGGAGATGTACTG 349

Qy     183 TGTCCGCTGCAAGCCTACAAAGTCAGCTCGTTCCATCCGGGCCCAGCGCCCACTGACAT 242
      |||
Db     348 TGCTCCGCTGAAGCCTACAAAGTCAGCTCGTTCCATCCGGGCCCAGCGCCCACTGACAT 289

Qy     243 GCCCAAGACTCAGAAGTCCCAGCCCCTATCGACACACAAGAAAAGGAAGCTGCAAAGGAG 302
      |||
Db     288 GCCCAAGACTCAGAAGTCCCAGCCCCTATCGACACACAAGAAAAGGAAGCTGCAAAGGAG 229

Qy     303 AAGGAAAGGAAGTACACTTGAAGAACACAAGTAGAGGAAGTGCAGGAAACAAGACCTACA 362
      |||
Db     228 AAGGAAAGGAAGTACACTTGAAGAACACAAGTAGAGGAAGTGCAGGAAACAAGACCTACA 169

Qy     363 GAATGTAGGAGGAGCCTCCCGAGGAACAGAAAATGCCACGTCACCGCAAGATCCTTTGCT 422
      |||
Db     168 GAATGTAGGAGGAGCCTCCCGAGGAACAGAAAATGCCACGTCACCGCAAGATCCTTTGCT 109

Qy     423 GCTTGAGCAACCTGCAAAACATCGGAACACCTGCCAAATATCAATAATGAGTTCAATATC 482
      |||
Db     108 GCTTGAGCAACCTGCAAAACATCGGAACACCTGCCAAATATCAATAATGAGTTCAATACC 49

Qy     483 ATTTTCAGAGATGGGCATTTCCCTCAATGAAATACACAAGTAAACATTC 530
      |||
  
```

Db

48 ATTTGAGAGATGGGCATTTCCCTCAATGAAATACACAAGTAAACATTTC 1

RESULT 6

AI265629/c

LOCUS

AI265629 558 bp mRNA linear EST 18-NOV-1998

DEFINITION

uj04b07.x1 Sugano mouse liver mlia Mus musculus cDNA clone  
IMAGE:1890901 3' similar to gb:X04482 Mouse mRNA for  
preproinsulin-like growth factor IB (MOUSE);, mRNA sequence.

ACCESSION

AI265629

VERSION

AI265629.1 GI:3883787

KEYWORDS

EST.

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 558)

AUTHORS

Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,  
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,  
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,  
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and  
Waterston,R.

TITLE

The WashU-HHMI Mouse EST Project

JOURNAL

Unpublished

COMMENT

Contact: Marra M/Mouse EST Project

WashU-HHMI Mouse EST Project

Washington University School of MedicineP

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LLNL ; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:975225

Seq primer: custom primer used

High quality sequence stop: 495.

FEATURES

Location/Qualifiers

source

1..558

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="C57BL"

/db\_xref="taxon:10090"

/clone="IMAGE:1890901"

/sex="female"

/dev\_stage="adult"

/lab\_host="DH10B"

/clone\_lib="Sugano mouse liver mlia"

/note="Organ: liver; Vector: pME18S-FL3; Site\_1: DraIII

(CACTGTGTG); Site\_2: DraIII (CACCATGTG); 1st strand cDNA

was primed with an oligo(dT) primer

[ATGTGGCCTTTTTTTTTTTTTTTTTT]; double-stranded cDNA was

ligated to a DraIII adaptor [TGTGGCCTACTGG], digested

and cloned into distinct DraIII sites of the pME18S-FL3

vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should

be used to isolate the cDNA insert. Size selection was

performed to exclude fragments <1.5kb. Library

constructed by Dr. Sumio Sugano (University of Tokyo

Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTCTGCTCTAAAAGCTGCG and 3' end primer CGACCTGCAGCTCGAGCACA."

BASE COUNT 106 a 135 c 156 g 161 t  
ORIGIN

Query Match 82.9%; Score 446.8; DB 9; Length 558;  
Best Local Similarity 92.7%; Pred. No. 2.6e-94;  
Matches 469; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

```

Qy      1 GGACCAGAGACCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTCGTGTGTGGACCA 60
      |||
Db      506 GGACCAGAGACCCTTTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGACCG 447

Qy      61 AGGGGCTTTTACTTCAACAAGCCACAGTCTATGGCTCCAGCATTTCGAGGGGCACCACAG 120
      |||
Db      446 AGGGGCTTTTACTTCAACAAGCCACAGGCTATGGCTCCAGCATTTCGAGGGGCACCTCAG 387

Qy      121 ACGGGCATTGTGGATGAGTGTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 180
      ||
Db      386 ACAGGCATTGTGGATGAGTGTGCTTCCGGAGCTGTGATCTGAGGAGACTGGAGATGTAC 327

Qy      181 TGTGTCCGCTGCAAGCCTACAAAGTCAGCTCGTTCCATCCGGGCCCAGCGCCCACTGAC 240
      ||||
Db      326 TGTGCCCCACTGAAGCCTACAAAAGCAGCCCGCTCTATCCGTGCCCAGCGCCCACTGAC 267

Qy      241 ATGCCCAAGACTCAGAAGTCCCAGCCCCTATCGACACACAAGAAAAGGAAGCTGCAAAGG 300
      |||
Db      266 ATGCCCAAGACTCAGAAGTCCCCGTCCCTATCGACAAAACAAGAAACGAAGCTGCAAAGG 207

Qy      301 AGAAGGAAAGGAAGTACACTTGAAGAACACAAGTAGAGGAAGTGCAGGAAACAAGACCTA 360
      |||
Db      206 AGAAGGAAAGGAAGTACATTTGAAGAACACAAGTAGAGGAAGTGCAGGAAACAAGACCTA 147

Qy      361 CAGAATGTAGGAGGAGCCTCCCGAGGAACAGAAAATGCCACGTCACCGCAAGATCCTTTG 420
      |||
Db      146 CAGAATGTAGGAGGAGCCTCCACGGAGCAGAAAATGCCACATCACCGCAGGATCCTTTG 87

Qy      421 CTGCTTGAGCAACCTGCAAAACATCGGAACACCTGCCAAATATCAATAATGAGTTCAATA 480
      |||
Db      86 CTGCTTGAGCAACCTGCAAAACATCGAAACACCTACCAAATAACAATAAAGTCCAATA 27

Qy      481 TCATTTTCAGAGATGGGCATTTCCCTC 506
      ||||
Db      26 ACATTACAAAGATGGGCATTTCCCCC 1

```

# RESULT 7

BQ200567/c

LOCUS BQ200567 653 bp mRNA linear EST 02-MAY-2002

DEFINITION UI-R-DZ1-cne-a-18-0-UI.s1 UI-R-DZ1 Rattus norvegicus cDNA clone  
UI-R-DZ1-cne-a-18-0-UI 3', mRNA sequence.

ACCESSION BQ200567

VERSION BQ200567.1 GI:20417032

KEYWORDS EST.

SOURCE Rattus norvegicus (Norway rat)

ORGANISM Rattus norvegicus



Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 653)

AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.

TITLE Normalization and subtraction: two approaches to facilitate gene discovery

JOURNAL Genome Res. 6 (9), 791-806 (1996)

MEDLINE 97044477

PUBMED 8889548

COMMENT Contact: Soares, MB  
 Coordinated Laboratory for Computational Genomics  
 University of Iowa  
 375 Newton Road , 4156 MEBRF, Iowa City, IA 52242, USA  
 Tel: 319 335 8250  
 Fax: 319 335 9565  
 Email: bento-soares@uiowa.edu  
 The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. The sequence tag present in the cDNA between the NotI site and the oligo-dT track served to verify it as a clone from the normalized cartilaginous tumor library cDNA Library Preparation: M.B. Soares Lab Clone distribution: clones will be available through Research Genetics (www.resgen.com) The following repetitive elements were found in this cDNA sequence: 1-43,  
 >AT\_rich#Low\_complexity 118-164, >POLY\_A#Simple\_repeat  
 Seq primer: M13 Forward  
 POLYA=Yes.

FEATURES

source Location/Qualifiers

1..653

/organism="Rattus norvegicus"

/mol\_type="mRNA"

/strain="Sprague-Dawley"

/db\_xref="taxon:10116"

/clone="UI-R-DZ1-cne-a-18-0-UI"

/tissue\_type="Chondrosarcoma"

/dev\_stage="37 days"

/lab\_host="DH10B (Life Technologies)"

/clone\_lib="UI-R-DZ1"

/note="Organ: Spine; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site\_1: Not I; Site\_2: Eco RI; UI-R-DZ1 is a normalized cDNA library containing the following tissue(s): Swarm Rat Chondrosarcoma. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is CATTCTTGTA. The Rat cartilaginous tumor tissue was provided by Dr Jeff Stevens at the University of Iowa. TAG\_LIB=UI-R-DZ1 TAG\_TISSUE=cartilaginous tumor"

TAG\_SEQ=CATTCTTGTA"  
BASE COUNT      134 a      137 c      137 g      245 t  
ORIGIN

Query Match                  81.1%;  Score 437;  DB 13;  Length 653;  
Best Local Similarity      97.6%;  Pred. No. 5.3e-92;  
Matches 454;  Conservative   0;  Mismatches 10;  Indels   1;  Gaps   1;

```
Qy      67 TTTTACTTCAACAAGCCACAGTCTATGGCTCCAGCATTCCGAGGGCACCACAGACGGGC 126
      |||
Db      653 TTTTACTTCAACAAGCCACAGGCTATGGCTCCAGCATTCCGAGGGCACCACAGACGGGC 594

Qy     127 ATTGTGGATGAGTGTGTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTACTGTGTC 186
      |||
Db     593 ATTGTGGATGAGTG-TGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTACTGTGCT 535

Qy     187 CGCTGCAAGCCTACAAAGTCAGCTCGTTCCATCCGGGCCCAGCGCCCACTGACATGCCC 246
      |
Db     534 CCGCTGAAGCCTACAAAGTCAGCTCGTTCCATCCGGGCCCAGCGCCCACTGACATGCCC 475

Qy     247 AAGACTCAGAAGTCCCAGCCCCCTATCGACACACAAGAAAAGGAAGCTGCAAAGGAGAAGG 306
      |||
Db     474 AAGACTCAGAAGTCCCAGCCCCCTATCGACACACAAGAAAAGGAAGCTGCAAAGGAGAAGG 415

Qy     307 AAAGGAAGTACACTTGAAGAACACAAGTAGAGGAAGTGCAGGAAACAAGACCTACAGAAT 366
      |
Db     414 AAAGGAAGTACACTTGAAGAACACAAGTAGAGGAAGTGCAGGAAACAAGACCTACAGAAT 355

Qy     367 GTAGGAGGAGCCTCCCGAGGAACAGAAAATGCCACGTCACCGCAAGATCCTTTGCTGCTT 426
      |||
Db     354 GTAGGAGGAGCCTCCCGAGGAACAGAAAATGCCACGTCACCGCAAGATCCTTTGTTGCTT 295

Qy     427 GAGCAACCTGCAAAACATCGGAACACCTGCCAAATATCAATAATGAGTTCAATATCATTT 486
      |||
Db     294 GAGCAACCTGCAAAACATCGGAACACCTGCCAAATATCAATAATGAGTTCAATACCATTT 235

Qy     487 CAGAGATGGGCATTTCCCTCAATGAAATACACAAGTAAACATTCC 531
      |||
Db     234 CAGAGATGGGCATTTCCCTCAATGAAATACACAAGTAAACATTCC 190
```

RESULT 8

BF383724

LOCUS          BF383724                  594 bp      mRNA      linear      EST 27-NOV-2000

DEFINITION     602044632F1 NCI\_CGAP\_Li9 Mus musculus cDNA clone IMAGE:4194295 5',  
                 mRNA sequence.

ACCESSION      BF383724

VERSION         BF383724.1  GI:11365029

KEYWORDS        EST.

SOURCE         Mus musculus (house mouse)

ORGANISM        Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE      1  (bases 1 to 594)

AUTHORS         NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE           National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL         Unpublished

COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-r@mail.nih.gov  
 Tissue Procurement: Jeffrey E. Green, M.D.  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LLAM9527 row: p column: 08  
 High quality sequence stop: 589.

FEATURES Location/Qualifiers  
 source 1. .594  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="FVB/N"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:4194295"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /clone\_lib="NCI\_CGAP\_Li9"  
 /note="Organ: liver; Vector: pCMV-SPORT6; Site\_1: NotI;  
 Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
 Average insert size 1.9 kb. Constructed by Life  
 Technologies. Note: this is a NCI\_CGAP Library."

BASE COUNT 175 a 162 c 142 g 115 t  
 ORIGIN

Query Match 80.1%; Score 431.8; DB 10; Length 594;  
 Best Local Similarity 92.4%; Pred. No. 8.5e-91;  
 Matches 465; Conservative 0; Mismatches 37; Indels 1; Gaps 1;

```

Qy      1 GGACCAGAGACCCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTCGTGTGTGGACCA 60
        |||
Db      93 GGACCAGAGACCC-TTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGACCG 151

Qy      61 AGGGGCTTTTACTTCAACAAGCCACAGTCTATGGCTCCAGCATTCCGAGGGGCACACAG 120
        |||
Db     152 AGGGGCTTTTACTTCAACAAGCCACAGGCTATGGCTCCAGCATTCCGAGGGGCACCTCAG 211

Qy     121 ACGGGCATTGTGGATGAGTGTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 180
        ||
Db     212 ACAGGCATTGTGGATGAGTGTGCTTCCGGAGCTGTGATCTGAGGAGACTGGAGATGTAC 271

Qy     181 TGTGTCCGCTGCAAGCCTACAAAGTCAGCTCGTTCCATCCGGGGCCAGCGCCCACTGAC 240
        ||||
Db     272 TGTGCCCCACTGAAGCCTACAAAAGCAGCCCGCTCTATCCGTGCCCAGCGCCCACTGAC 331

Qy     241 ATGCCCAAGACTCAGAAGTCCCAGCCCCTATCGACACACAAGAAAAGGAAGCTGCAAAGG 300
        |||
Db     332 ATGCCCAAGACTCAGAAGTCCCCGTCCCTATCGACAAACAAGAAAACGAAGCTGCAAAGG 391

Qy     301 AGAAGGAAAGGAAGTACACTTGAAGAACACAAGTAGAGGAAGTGCAGGAAACAAGACCTA 360
        |||
Db     392 AGAAGGAAAGGAAGTACATTTGAAGAACACAAGTAGAGGAAGTGCAGGAAACAAGACCTA 451

Qy     361 CAGAAATGTAGGAGGAGCCTCCCGAGGAACAGAAAATGCCACGTCACCGCAAGATCCTTTG 420
        |||

```



double-stranded cDNA was ligated to a DraIII adaptor [TGTTGGCCTACTGG], digested and cloned into distinct DraIII sites of the pME18S-FL3 vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTCTGCTCTAAAAGCTGCG and 3' end primer CGACCTGCAGCTCGAGCACA."

BASE COUNT        230 a        219 c        172 g        187 t        8 others  
ORIGIN

Query Match                78.9%;    Score 425.2;    DB 9;    Length 816;  
Best Local Similarity    90.7%;    Pred. No. 3.2e-89;  
Matches 448;    Conservative    0;    Mismatches    46;    Indels        0;    Gaps        0;

```

Qy      1 GGACCAGAGACCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTCGTGTGTGGACCA 60
      |||
Db    323 GGACCAGAGACCCTTTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGACCG 382

Qy      61 AGGGGCTTTTACTTCAACAAGCCACAGTCTATGGCTCCAGCATTCCGAGGGGCACCACAG 120
      |||
Db    383 AGGGGCTTTTACTTCAACAAGCCACAGGCTATGGCTCCAGCATTCCGAGGGGCACCTCAG 442

Qy     121 ACGGGCATTGTGGATGAGTGTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 180
      |||
Db    443 ACAGGCATTGTGGATGAGTGTGCTTCCGGAGCTGTGATCTGAGGAGACTGGAGATGTAC 502

Qy     181 TGTGTCCGCTGCAAGCCTACAAAGTCAGCTCGTTCCATCCGGGCCCAGCGCCCACTGAC 240
      |||||
Db    503 TGTGCCCCACTGAAGCCTACAAAAGCAGCCGCTCTATCCGTGCCAGCGCCCACTGAC 562

Qy     241 ATGCCCAAGACTCAGAAGTCCCAGCCCCTATCGACACACAAGAAAAGGAAGCTGCAAAGG 300
      |||||
Db    563 ATGCCCAAGACTCAGAAGTCCCCGTCCCTATCGACAAACAAGAAAACGAAGCTGCAAAGG 622

Qy     301 AGAAGGAAAGGAAGTACACTTGAAGAACACAAGTAGAGGAAGTGCAGGAAACAAGACCTA 360
      |||||
Db    623 AGAAGGAAAGGAAGTACATTGAAGAACACAAGTAGAGGAAGTGCANGAAACAAGACCTA 682

Qy     361 CAGAATGTAGGAGGAGCCTCCCGAGGAACAGAAAATGCCACGTCACCGCAAGATCCTTTG 420
      |||||
Db    683 CAGAATGTANGAGGAGCCTNCCACGGAGCAGAAATGCCACATCACCGCANGATCCTTTG 742

Qy     421 CTGCTTGAGCAACCTGCAAAACATCGGAACACCTGCCAAATATCAATAATGAGTTCAATA 480
      |||||
Db    743 CTGCTTGAGCAACCTGCANAACATCGAAACACCTACCAAATAACATNTATAAGTCCAATN 802

Qy     481 TCATTTTCAGAGATG 494
      |||||
Db    803 ACATTACCAAGATG 816

```

RESULT 10  
AW495481/c  
LOCUS

AW495481

499 bp

mRNA

linear

EST 24-FEB-2000

DEFINITION UI-M-BH3-ayy-g-11-0-UI.s1 NIH\_BMAP\_M\_S4 Mus musculus cDNA clone  
 UI-M-BH3-ayy-g-11-0-UI 3', mRNA sequence.

ACCESSION AW495481

VERSION AW495481.1 GI:7065762

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 499)

AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.

TITLE Normalization and subtraction: two approaches to facilitate gene  
 discovery

JOURNAL Genome Res. 6 (9), 791-806 (1996)

MEDLINE 97044477

PUBMED 8889548

COMMENT Contact: Chin, H  
 National Institute of Mental Health  
 6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD  
 20892-9643, USA  
 Tel: 301 443 1706  
 Fax: 301 443 9890  
 Email: mEST@mail.nih.gov  
 The sequence contained an oligo-dT track that was present in the  
 oligonucleotide that was used to prime the synthesis of first  
 strand cDNA and therefore this may represent a bonafide poly A  
 tail. The sequence tag present in the cDNA between the NotI site  
 and the oligo-dT track served to identify it as a clone from the  
 normalized pineal glands library cDNA Library Preparation: M.B.  
 Soares Lab Clone distribution: Researchers may obtain BMAP cDNA  
 clones from RESEARCH GENETICS. It should be noted that Bento Soares  
 is generating a small number of additional specialized  
 non-redundant arrays of BMAP cDNAs whose availability will be  
 considered under appropriate and limited collaborative arrangements  
 Seq primer: M13 Forward  
 POLYA=Yes.

FEATURES Location/Qualifiers

source 1..499  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="UI-M-BH3-ayy-g-11-0-UI"  
 /dev\_stage="27-32 days"  
 /lab\_host="DH10B (Life Technologies)"  
 /clone\_lib="NIH\_BMAP\_M\_S4"  
 /note="Vector: pT7T3D-Pac (Pharmacia) with a modified  
 polylinker; Site\_1: Not I; Site\_2: Eco RI; The  
 NIH\_BMAP\_M\_S4 library is a subtracted library of a series,  
 ultimately derived from a mixture of individually tagged  
 normalized libraries from ten regions of the mouse brain  
 (cerebellum, brain stems, olfactory bulbs, hypothalamus,  
 cortex, amygdala, basal ganglia, pineal gland, striatum,  
 hippocampus) after a series of subtractions to reduce the  
 representation of cDNAs from which ESTs had already been  
 generated. The following serially subtracted libraries  
 were generated in this process: NIH\_BMAP\_M\_S4,

NIH\_BMAP\_M\_S3.3, NIH\_BMAP\_M\_S3.2, NIH\_BMAP\_M\_S3.1,  
 NIH\_BMAP\_M\_S2, NIH\_BMAP\_M\_S1. The subtracted library  
 (NIH\_BMAP\_M\_S4) was constructed as follows: PCR amplified  
 cDNA inserts from NIH\_BMAP\_M\_S3.3, NIH\_BMAP\_M\_S3.2, and  
 NIH\_BMAP\_M\_S3.1 clones from which 3' ESTs had been derived  
 was used as a driver in a hybridization with a pool of  
 the NIH\_BMAP\_M\_S3.3, NIH\_BMAP\_M\_S3.2, and NIH\_BMAP\_M\_S3.1  
 libraries in the form of single-stranded circles. The  
 remaining single-stranded circles (subtracted library)  
 was purified by hydroxyapatite column chromatography,  
 converted to double-stranded circles and electroporated  
 into DH10B bacteria (Life Technologies) to generate the  
 NIH\_BMAP\_M\_S4 library. This procedure has been previously  
 described (Bonaldo, Lennon and Soares, Genome Research  
 6:791-806, 1996)  
 TAG\_LIB=NIH\_BMAP\_M\_S4  
 TAG\_TISSUE=pineal-glands  
 TAG\_SEQ=CAGAC"

BASE COUNT        86 a     112 c     124 g     177 t  
 ORIGIN

Query Match                    78.5%;    Score 423;    DB 9;    Length 499;  
 Best Local Similarity        91.8%;    Pred. No. 9.5e-89;  
 Matches 447;    Conservative        0;    Mismatches    40;    Indels        0;    Gaps        0;

```

Qy      50 TGTGTGGACCAAGGGGCTTTTACTTCAACAAGCCACAGTCTATGGCTCCAGCATTCGGA 109
      |||||
Db     499 TGTGTGGACCGAGGGGCTTTTACTTCAACAAGCCACAGGCTATGGCTCCAGCATTCGGA 440

Qy     110 GGGCACCACAGACGGGCATTGTGGATGAGTGTGCTTCCGGAGCTGTGATCTGAGGAGGC 169
      |||||
Db     439 GGGCACCTCAGACAGGCATTGTGGATGAGTGTGCTTCCGGAGCTGTGATCTGAGGAGAC 380

Qy     170 TGGAGATGTACTGTGTCCGCTGCAAGCCTACAAAGTCAGCTCGTTCCATCCGGGCCCAGC 229
      |||||
Db     379 TGGAGATGTACTGTGCCCCACTGAAGCCTACAAAAGCAGCCCGCTCTATCCGTGCCCAGC 320

Qy     230 GCCCACTGACATGCCCAAGACTCAGAAGTCCCAGCCCCTATCGACACACAAGAAAAGGA 289
      |||||
Db     319 GCCCACTGACATGCCCAAGACTCAGAAGTCCCCGTCCTATCGACAAACAAGAAAACGA 260

Qy     290 AGCTGCAAAGGAGAAGGAAAGGAAGTACACTTGAAGAACACAAGTAGAGGAAGTGCAGGA 349
      |||||
Db     259 AGCTGCAAAGGAGAAGGAAAGGAAGTACATTTGAAGAACACAAGTAGAGGAAGTGCAGGA 200

Qy     350 AACAAAGACCTACAGAATGTAGGAGGAGCCTCCCAGGGAACAGAAAATGCCACGTACCCGC 409
      |||||
Db     199 AACAAAGACCTACAGAATGTAGGAGGAGCCTCCCAGGAGCAGAAAATGCCACATACCCGC 140

Qy     410 AAGATCCTTTGCTGCTTGAGCAACCTGCAAAACATCGGAACACCTGCCAAATATCAATAA 469
      |||||
Db     139 AGGATCCTTTGCTGCTTGAGCAACCTGCAAAACATCGAAACACCTACCAAATAACAATAA 80

Qy     470 TGAGTTCAATATCATTTTCAGAGATGGGCATTTCCCTCAATGAAATACACAAGTAAACATT 529
      |||||
Db      79 TAAGTCCAATAACATTACAAAGATGGGCATTTCCCCCAATGAAATATACAAGTAAACATT 20

```

QY 530 CCCGGAA 536  
|| ||  
Db 19 CCAAAAA 13

RESULT 11

AI876493/c

LOCUS AI876493 642 bp mRNA linear EST 21-JUL-1999

DEFINITION uj59b10.x1 Sugano mouse liver mlia Mus musculus cDNA clone  
IMAGE:1924219 3' similar to gb:X57025\_rnal INSULIN-LIKE GROWTH  
FACTOR IA PRECURSOR (HUMAN); gb:X04482 Mouse mRNA for  
preproinsulin-like growth factor IB (MOUSE);, mRNA sequence.

ACCESSION AI876493

VERSION AI876493.1 GI:5550542

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 642)

AUTHORS Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,  
Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person  
,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter  
,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,  
Waterston,R. and Wilson,R.

TITLE The WashU-NCI Mouse EST Project 1999

JOURNAL Unpublished

COMMENT Contact: Marra M/WashU-NCI Mouse EST Project 1999  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: mouseest@watson.wustl.edu  
This clone is available royalty-free through LLNL ; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
MGI:980511

Seq primer: custom primer used  
High quality sequence stop: 257.

FEATURES

source

Location/Qualifiers

1: .642

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="C57BL"

/db\_xref="taxon:10090"

/clone="IMAGE:1924219"

/sex="female"

/dev\_stage="adult"

/lab\_host="DH10B"

/clone\_lib="Sugano mouse liver mlia"

/note="Organ: liver; Vector: pME18S-FL3; Site\_1: DraIII  
(CACTGTGTG); Site\_2: DraIII (CACCATGTG); 1st strand cDNA  
was primed with an oligo(dT) primer

[ATGTGGCCTTTTTTTTTTTTTTTTTT]; double-stranded cDNA was  
ligated to a DraIII adaptor [TGTGGCCTACTGG], digested  
and cloned into distinct DraIII sites of the pME18S-FL3  
vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should  
be used to isolate the cDNA insert. Size selection was



performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTCTGCTCTAAAAGCTGCG and 3' end primer CGACCTGCAGCTCGAGCACA."

BASE COUNT        127 a        154 c        175 g        185 t        1 others  
ORIGIN

Query Match                    78.1%;    Score 421.2;    DB 9;    Length 642;  
Best Local Similarity        91.1%;    Pred. No. 2.6e-88;  
Matches 458;    Conservative        0;    Mismatches 44;    Indels        1;    Gaps        1;

```

Qy      2  GACCAGAGACCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTCGTGTGTGGACCAA 61
      |||
Db    503  GACCAGAGACCCTTTGCGGGGCTGAGCTGGTGGATGCTCTTCAGGTCTGTGTGTGGACCGA 444
      |||

Qy      62  GGGGCTTTTACTTCAACAAGCCCACAGTCTATGGCTCCAGCATTCGGAGGGCACCACAGA 121
      |||
Db    443  GGGGCTTTTCTTCAACAAGGCCACAGGCTATGGCTCCAGCATTTGGAGGGCACCTCAGA 384
      |||

Qy     122  CGGGCATTGTGGATGAGTGTGCTTCCGG-AGCTGTGATCTGAGGAGGCTGGAGATGTAC 180
      |||
Db    383  CAGTCAATGTGGATGAGTGTGCTTCCGGAAGCTGTGATCTGAGAAGACTGNAGATGTAC 324
      |||

Qy     181  TGTGTCCGCTGCAAGCCTACAAAGTCAGCTCGTTCCATCCGGGCCCAGCGCCCACTGAC 240
      |||
Db    323  TGTGCCCCACTGAAGCCTACAAAAGCAGCCGCTCTATCCGTGCCCAGCGCCCACTGAC 264
      |||

Qy     241  ATGCCCAAGACTCAGAAGTCCCAGCCCCATCGACACACAAGAAAAGGAAGCTGCAAAGG 300
      |||
Db    263  ATGCCCAAGACTCAGAAGTCCCCGTCCCTATCGACAAACAAGAAACGAAGCTGCAAAGG 204
      |||

Qy     301  AGAAGGAAAGGAAGTACACTTGAAGAACACAAGTAGAGGAAGTGCAGGAAACAAGACCTA 360
      |||
Db    203  AGAAGGAAAGGAAGTACATTGAAGAACACAAGTAGAGGAAGTGCAGGAAACAAGACCTA 144
      |||

Qy     361  CAGAATGTAGGAGGAGCCTCCCGAGGAACAGAAAATGCCACGTACCGCAAGATCCTTTG 420
      |||
Db    143  CAGAATGTAGGAGGAGCCTCCCGAGGAGCAGAAAATGCCACATCACCGCAGGATCCTTTG 84
      |||

Qy     421  CTGCTTGAGCAACCTGCAAAACATCGGAACACCTGCCAAATATCAATAATGAGTTCAATA 480
      |||
Db     83  CTGCTTGAGCAACCTGCAAAACATCGAAACACCTACCAAATAACAATAAAGTCCAATA 24
      |||

Qy     481  TCATTTTCAGAGATGGGCATTTCC 503
      |||
Db     23  ACATTACAAAGATGGGCATTTCC 1

```

# RESULT 12

AA963258/c

LOCUS            AA963258                    525 bp        mRNA        linear        EST 09-MAR-1999

DEFINITION      UI-R-E1-gh-f-04-0-UI.s1 UI-R-E1 Rattus norvegicus cDNA clone

UI-R-E1-gh-f-04-0-UI 3', mRNA sequence.

ACCESSION        AA963258

VERSION          AA963258.1    GI:4278182

KEYWORDS        EST.

SOURCE Rattus norvegicus (Norway rat)

ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

REFERENCE 1 (bases 1 to 525)

AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.

TITLE Normalization and subtraction: two approaches to facilitate gene  
discovery

JOURNAL Genome Res. 6 (9), 791-806 (1996)

MEDLINE 97044477

PUBMED 8889548

COMMENT On May 18, 1998 this sequence version replaced gi:3136750.  
Contact: Soares, MB  
Coordinated Laboratory for Computational Genomics  
University of Iowa  
375 Newton Road , 4156 MEBRF, Iowa City, IA 52242, USA  
Tel: 319 335 8250  
Fax: 319 335 9565  
Email: bento-soares@uiowa.edu  
The sequence tag present in the cDNA between the NotI site and the  
oligo-dT track served to identify it as a clone from the normalized  
adult 12-Day-Embryo library. cDNA Library Preparation: M. Fatima  
Bonaldo, Ph.D. Clone distribution: clones will be available through  
Research Genetics  
Seq primer: M13 Forward.

FEATURES

source Location/Qualifiers

1. .525  
/organism="Rattus norvegicus"  
/mol\_type="mRNA"  
/strain="Sprague-Dawley"  
/db\_xref="taxon:10116"  
/clone="UI-R-E1-gh-f-04-0-UI"  
/dev\_stage="adult"  
/lab\_host="DH10B (Life Technologies)"  
/clone\_lib="UI-R-E1"  
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified  
polylinker; Site\_1: Not I; Site\_2: Eco RI; The UI-R-E1  
library is a subtracted library derived from the UI-R-E0  
library. The UI-R-E0 library consisted of a mixture of  
individually tagged normalized libraries constructed from  
8, 12 and 18-day embryo. The tag is a string of 3-5  
nucleotides present between the Not I site and the  
oligo-dT track which allows identification of the library  
of origin of a clone within the mixture. The subtracted  
library (UI-R-E1) was constructed as follows: PCR  
amplified cDNA inserts from a pool of UI-R-E0 clones from  
which 3' ESTs had been derived was used as a driver in a  
hybridization with the UI-R-E0 library in the form of  
single-stranded circles. The remaining single-stranded  
circles (subtracted library) was purified by  
hydroxyapatite column chromatography, converted to  
double-stranded circles and electroporated into DH10B  
bacteria (Life Technologies) to generate the UI-R-E1  
library. This procedure has been previously described  
(Bonaldo, Lennon and Soares, Genome Research 6: 791-806,  
1996)"

BASE COUNT        107 a        133 c        126 g        159 t  
ORIGIN

Query Match                74.7%;    Score 402.6;    DB 9;    Length 525;  
Best Local Similarity    88.5%;    Pred. No. 5.8e-84;  
Matches 470;    Conservative    0;    Mismatches    9;    Indels    52;    Gaps    1;

```
Qy      1 GGACCAGAGACCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTCGTGTGTGGACCA 60
      |||
Db     521 GGACCAGAGACCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTCGTGTGTGGACCA 462

Qy      61 AGGGGCTTTTACTTCAACAAGCCCAAGTCTATGGCTCCAGCATTTCGGAGGGCACCACAG 120
      |||
Db     461 AGGGGCTTTTACTTCAACAAGCCCAAGTCTATGGCTCCAGCATTTCGGAGGGCACCACAG 402

Qy     121 ACGGGCATTGTGGATGAGTGTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 180
      |||
Db     401 ACGGGCATTGTGGATGAGTGTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 342

Qy     181 TGTGTCCGCTGCAAGCCTACAAAGTCAGCTCGTTCCATCCGGGCCCAGCGCCCACTGAC 240
      |||
Db     341 TGTGTCCGCTGCAAGCCTACAAAGTCAGCTCGTTCCATCCGGGCCCAGCGCCCACTGAC 282

Qy     241 ATGCCCAAGACTCAGAAGTCCCAGCCCCTATCGACACACAAGAAAAGGAAGCTGCAAAGG 300
      |||
Db     281 ATGCCCAAGACTCAG----- 267

Qy     301 AGAAGGAAAGGAAGTACACTTGAAGAACACAAGTAGAGGAAGTGCAGGAAACAAGACCTA 360
      |||
Db     266 -----AAGGAAGTACACTTGAAGAACACAAGTAGAGGAAGTGCAGGAAACAAGACCTA 214

Qy     361 CAGAATGTAGGAGGAGCCTCCCGAGGAACAGAAAATGCCACGTACCGCAAGATCCTTTG 420
      |||
Db     213 CAGAATGTAGGAGGAGCCTCCCGAGGAACAGAAAATGCCACGTACCGCAAGATCCTTTG 154

Qy     421 CTGCTTGAGCAACCTGCAAAACATCGGAACACCTGCCAAATATCAATAATGAGTTCAATA 480
      |||
Db     153 CTGCTTGAGCAACCTGCAAAACATCGGAACACCTGCCAAATATCAATAATGAGTTCAATA 94

Qy     481 TCATTTTCAGAGATGGGCATTTCCCTCAATGAAATACACAAGTAAACATTCC 531
      |||
Db      93 CCATTTTCAGAGATGGGCATTTCCCTCAATGAAATACACAAGTAAACATTCC 43
```

RESULT 13

AA945553/c

LOCUS        AA945553                500 bp        mRNA        linear        EST 08-JAN-1999

DEFINITION   EST201052 Normalized rat liver, Bento Soares Rattus sp. cDNA clone  
              RLIAO83 3' end, mRNA sequence.

ACCESSION    AA945553

VERSION      AA945553.1    GI:4132547

KEYWORDS     EST.

SOURCE      Rattus sp.

ORGANISM     Rattus sp.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

REFERENCE 1 (bases 1 to 500)  
AUTHORS Lee,N.H., Glodek,A., Chandra,I., Mason,T.M., Quackenbush,J.,  
Kerlavage,A.R. and Adams,M.D.  
TITLE Rat Genome Project: Generation of a Rat EST (REST) Catalog & Rat  
Gene Index  
JOURNAL Unpublished  
COMMENT On May 1, 1998 this sequence version replaced gi:3105469.  
Contact: Lee, NH  
The Institute for Genomic Research  
9712, Medical Center Drive, Rockville, MD 20850, USA  
Tel: (301)-838-3529  
Fax: (301)-838-0208  
Email: nhlee@tigr.org  
Seq primer: M13-21.

FEATURES Location/Qualifiers  
source 1..500  
/organism="Rattus sp."  
/mol\_type="mRNA"  
/db\_xref="taxon:10118"  
/clone="RLIA083"  
/clone\_lib="Normalized rat liver, Bento Soares"  
/note="Organ: liver; Vector: pT7T3Pac; Site\_1: EcoRI;  
Site\_2: NotI"

BASE COUNT 99 a 129 c 130 g 142 t  
ORIGIN

Query Match 74.5%; Score 401.6; DB 9; Length 500;  
Best Local Similarity 88.5%; Pred. No. 9.8e-84;  
Matches 469; Conservative 0; Mismatches 9; Indels 52; Gaps 1;

```

Qy      1 GGACCAGAGACCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTCGTGTGTGGACCA 60
      |||
Db      478 GGACCAGAGACCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTCGTGTGTGGACCA 419
      |||

Qy      61 AGGGGCTTTTACTTCAACAAGCCCACAGTCTATGGCTCCAGCATTCCGAGGGGCACCAACAG 120
      |||
Db      418 AGGGGCTTTTACTTCAACAAGCCCACAGGCTATGGCTCCAGCATTCCGAGGGGCACCAACAG 359
      |||

Qy      121 ACGGGCATTGTGGATGAGTGTGTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 180
      |||
Db      358 ACGGGCATTGTGGATGAGTGTGTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 299
      |||

Qy      181 TGTGTCCGCTGCAAGCCTACAAAGTCAGCTCGTTCCATCCGGGGCCAGCGCCCACTGAC 240
      |||
Db      298 TGTGTCCGCTGCAAGCCTACAAAGTCAGCTCGTTCCATCCGGGGCCAGCGCCCACTGAC 239
      |||

Qy      241 ATGCCCAAGACTCAGAAGTCCCAGCCCCTATCGACACACAAGAAAAGGAAGCTGCAAAGG 300
      |||
Db      238 ATGCCCAAGACTCAG----- 224

Qy      301 AGAAGGAAAGGAAGTACACTTTGAAGAACACAAGTAGAGGAAGTGCAGGAAACAAGACCTA 360
      |||
Db      223 -----AAGGAAGTACACTTTGAAGAACACAAGTAGAGGAAGTGCAGGAAACAAGACCTA 171
      |||

Qy      361 CAGAAATGTAGGAGGAGCCTCCCGAGGAACAGAAAATGCCACGTCACCGCAAGATCCTTTG 420
      |||
Db      170 CAGAAATGTAGGAGGAGCCTCCCGAGGAACAGAAAATGCCACGTCACCGCAAGATCCTTTG 111
      |||

```

Qy 421 CTGCTTGAGCAACCTGCAAAACATCGGAACACCTGCCAAATATCAATAATGAGTTCAATA 480  
 |||  
 Db 110 CTGCTTGAGCAACCTGCAAAACATCGGAACACCTGCCAAATATCAATAATGAGTTCAATA 51  
 Qy 481 TCATTTTCAGAGATGGGCATTTCCCTCAATGAAATACACAAGTAAACATTTC 530  
 |||  
 Db 50 CCATTTTCAGAGATGGGCATTTCCCTCAATGAAATACACAAGTAAACATTTC 1

RESULT 14

AI599751/c

LOCUS AI599751 525 bp mRNA linear EST 21-APR-1999

DEFINITION EST251454 Normalized rat embryo, Bento Soares Rattus sp. cDNA clone  
 REMEG86 3' end, mRNA sequence.

ACCESSION AI599751

VERSION AI599751.1 GI:4608799

KEYWORDS EST.

SOURCE Rattus sp.

ORGANISM Rattus sp.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.

REFERENCE 1 (bases 1 to 525)

AUTHORS Lee,N.H., Glodek,A., Chandra,I., Mason,T.M., Quackenbush,J.,  
 Kerlavage,A.R. and Adams,M.D.

TITLE Rat Genome Project: Generation of a Rat EST (REST) Catalog & Rat  
 Gene Index

JOURNAL Unpublished

COMMENT Contact: Lee, NH

The Institute for Genomic Research  
 9712, Medical Center Drive, Rockville, MD 20850, USA  
 Tel: (301)-838-3529  
 Fax: (301)-838-0208  
 Email: nhlee@tigr.org  
 Seq primer: M13-21.

FEATURES Location/Qualifiers

source 1. 525  
 /organism="Rattus sp."  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10118"  
 /clone="REMEG86"  
 /dev\_stage="embryo 8, 12, 18 dpc"  
 /clone\_lib="Normalized rat embryo, Bento Soares"  
 /note="Vector: pT7T3Pac; Site\_1: EcoRI; Site\_2: NotI"

BASE COUNT 107 a 133 c 140 g 145 t

ORIGIN

Query Match 74.1%; Score 399.4; DB 9; Length 525;  
 Best Local Similarity 88.1%; Pred. No. 3.2e-83;  
 Matches 468; Conservative 0; Mismatches 11; Indels 52; Gaps 1;

Qy 1 GGACCAGAGACCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTTCGTGTGTGGACCA 60  
 |||  
 Db 480 GGACCAGAGACCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAATTCGTGTGTGGACCA 421  
 Qy 61 AGGGGCTTTTACTTCAACAAGCCACAGTCTATGGCTCCAGCATTCGGAGGGGCACCACAG 120

```

Db      420 AGGGGCTTTTACTTCAACAAGCCACAGGCTATGGCTCCAGCATTTCGGAGGGCACCACAG 361
Qy      121 ACGGGCATTGTGGATGAGTGTGTGCTTCCGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 180
Db      360 ACGGGCATTGTGGATGAGTGTGTGCTTCCGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 301
Qy      181 TGTGTCCGCTGCAAGCCTACAAAGTCAGCTCGTTCCATCCGGGCCCAGCGCCACACTGAC 240
Db      300 TGTGCTCCGCTGAAGCCTACAAAGTCAGCTCGTTCCATCCGGGCCCAGCGCCACATTGAC 241
Qy      241 ATGCCCAAGACTCAGAAGTCCCAGCCCCTATCGACACACAAGAAAAGGAAGCTGCAAAGG 300
Db      240 ATGCCCAAGACTCAG----- 226
Qy      301 AGAAGGAAAGGAAGTACACTTGAAGAACACAAGTAGAGGAAGTGCAGGAAACAAGACCTA 360
Db      225 -----AAGGAAGTACACTTGAAGAACACAAGTAGAGGAAGTGCAGGAAACAAGACCTA 173
Qy      361 CAGAATGTAGGAGGAGCCTCCCGAGGAACAGAAAATGCCACGTCACCGCAAGATCCTTTG 420
Db      172 CAGAATGTAGGAGGAGCCTCCCGAGGAACAGAAAATGCCACGTCACCGCAAGATCCTTTG 113
Qy      421 CTGCTTGAGCAACCTGCAAAACATCGGAACACCTGCCAAATATCAATAATGAGTTCAATA 480
Db      112 CTGCTTGAGCAACCTGCAAAACATCGGAACACCTGCCAAATATCAATAATGAGTTCAATA 53
Qy      481 TCATTTTCAGAGATGGGCATTTCCCTCAATGAAATACACAAGTAAACATTCC 531
Db      52 CCATTTTCAGAGATGGGCATTTCCCTCAATGAAATACACAAGTAAACATTCC 2

```

```

RESULT 15
BI294072/c
LOCUS      BI294072              499 bp      mRNA      linear      EST 19-JUL-2001
DEFINITION UI-R-DK0-cej-b-03-0-UI.s1 UI-R-DK0 Rattus norvegicus cDNA clone
            UI-R-DK0-cej-b-03-0-UI 3', mRNA sequence.
ACCESSION  BI294072
VERSION    BI294072.1  GI:14956179
KEYWORDS   EST.
SOURCE     Rattus norvegicus (Norway rat)
ORGANISM   Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE  1 (bases 1 to 499)
AUTHORS    Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE      Normalization and subtraction: two approaches to facilitate gene
            discovery
JOURNAL    Genome Res. 6 (9), 791-806 (1996)
MEDLINE    97044477
PUBMED     8889548
COMMENT    Contact: Soares, MB
            Coordinated Laboratory for Computational Genomics
            University of Iowa
            375 Newton Road , 4156 MEBRF, Iowa City, IA 52242, USA
            Tel: 319 335 8250

```

Fax: 319 335 9565

Email: bento-soares@uiowa.edu

The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. The sequence tag present in the cDNA between the NotI site and the oligo-dT track served to identify it as a clone from the normalized rat heart pool library cDNA Library Preparation: M.B. Soares Lab Clone distribution: clones will be available through Research Genetics (www.resgen.com)

Seq primer: M13 Forward

POLYA=Yes.

#### FEATURES

source

#### Location/Qualifiers

1. .499

/organism="Rattus norvegicus"

/mol\_type="mRNA"

/strain="Sprague-Dawley"

/db\_xref="taxon:10116"

/clone="UI-R-DK0-cej-b-03-0-UI"

/dev\_stage="ADULT"

/lab\_host="DH10B (Life Technologies)"

/clone\_lib="UI-R-DK0"

/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site\_1: Not I; Site\_2: Eco RI; The UI-R-DK0 library is a subtracted library derived from a mixture of five individually tagged normalized rat libraries: brain-nRBP (20%), heart-nRHP (20%), kidney-nRKP (20%), aorta-nRAP (20%), and placenta-nRPP (20%). Each original library was constructed from a mixture of equal amounts of RNA from seven different developmental time-points: embryonic day 17, embryonic day 19, embryonic day 21, adult day 1, adult day 12, adult day 75, and adult day 200. (Exception: the aorta pool does not contain embryonic day 17 RNA and the placenta pool contains only the three embryonic stages). Each library was normalized individually according to the procedure described by Bonaldo, Lennon & Soares (Genome Research Genome 6: 791-806, 1996). For construction of the DK0 subtracted library, plasmid DNA from each of the five individually tagged normalized libraries was mixed in the proportions specified above and electroporated into competent bacteria for production of single-stranded circular DNA representing the pool of libraries. Single-stranded circular DNA representing these five normalized libraries was then used as a tracer in a subtractive hybridization with a driver (PCR amplified inserts from a plasmid DNA template preparation) comprising: a) a set of about 1,000 arrayed clones from each of the five non-normalized libraries of brain (CT0s), heart (CS0s), kidney (CU0s), aorta (CW0s), and placenta (CX0s). The resulting pool of approximately 5,000 clones represented about 33.3% of the final driver population. A set of about 2,000 arrayed clones from each of the five normalized libraries of brain (CT0), heart (CS0), kidney (CU0), aorta (CW0), and placenta (CX0). The resulting pool of about 10,000 clones represented about 66.6% of the final driver population.

TAG\_LIB=UI-R-DK0

TAG\_TISSUE=rat heart pool  
 TAG\_SEQ=ATAAGATAAC"  
 BASE COUNT      97 a      124 c      121 g      156 t      1 others  
 ORIGIN

Query Match                      73.3%;    Score 395;    DB 12;    Length 499;  
 Best Local Similarity    88.0%;    Pred. No. 3.4e-82;  
 Matches 463;    Conservative    0;    Mismatches    11;    Indels    52;    Gaps    1;

```

Qy      6 AGAGACCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTCGTGTGTGGACCAAGGGG 65
      |||
Db     499 AGAGACCCTNTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTCGTGTGTGGACCAAGGGG 440

Qy     66 CTTTTACTTCAACAAGCCCACAGTCTATGGCTCCAGCATTTCGAGGGGCACCACAGACGGG 125
      |||
Db    439 CTTTTACTTCAACAAGCCCACAGGCTATGGCTCCAGCATTTCGAGGGGCACCACAGACGGG 380

Qy    126 CATTGTGGATGAGTGTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTACTGTGT 185
      |||
Db    379 CATTGTGGATGAGTGTGCTTCCGGAGCTGTTATCTGAGGAGGCTGGAGATGTACTGTGC 320

Qy    186 CCGCTGCAAGCCTACAAAGTCAGCTCGTTCCATCCGGGGCCAGCGCCCACTGACATGCC 245
      |||
Db    319 TCCGCTGAAGCCTACAAAGTCAGCTCGTTCCATCCGGGGCCAGCGCCCACTGACATGCC 260

Qy    246 CAAGACTCAGAAGTCCCAGCCCCTATCGACACACAAGAAAAGGAAGCTGCAAAGGAGAAG 305
      |||
Db    259 CAAGACTCAG----- 250

Qy    306 GAAAGGAAGTACACTTGAAGAACACAAGTAGAGGAAGTGACAGGAAACAAGACCTACAGAA 365
      |||
Db    249 --AAGGAAGTACACTTGAAGAACACAAGTAGAGGAAGTGACAGGAAACAAGACCTACAGAA 192

Qy    366 TGTAGGAGGAGCCTCCCGAGGAACAGAAAATGCCACGTCACCGCAAGATCCTTTGCTGCT 425
      |||
Db    191 TGTAGGAGGAGCCTCCCGAGGAACAGAAAATGCCACGTCACCGCAAGATCCTTTGCTGCT 132

Qy    426 TGAGCAACCTGCAAAACATCGGAACACCTGCCAAATATCAATAATGAGTTCAATATCATT 485
      |||
Db    131 TGAGCAACCTGCAAAACATCGGAACACCTGCCAAATATCAATAATGAGTTCAATACCATT 72

Qy    486 TCAGAGATGGGCATTTCCCTCAATGAAATACACAAGTAAACATTCC 531
      |||
Db    71 TCAGAGATGGGCATTTCCCTCAATGAAATACACAAGTAAACATTCC 26
  
```

Search completed: December 13, 2003, 07:29:50  
 Job time : 1758.93 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 13, 2003, 05:41:20 ; Search time 2408.26 Seconds  
(without alignments)  
9156.102 Million cell updates/sec

Title: US-09-852-261-3  
Perfect score: 539  
Sequence: 1 ggaccagagaccctttgcgg.....agtaaacattccccggaattc 539

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vi:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*

28: em\_un:\*  
 29: em\_vi:\*  
 30: em\_htg\_hum:\*  
 31: em\_htg\_inv:\*  
 32: em\_htg\_other:\*  
 33: em\_htg\_mus:\*  
 34: em\_htg\_pln:\*  
 35: em\_htg\_rod:\*  
 36: em\_htg\_mam:\*  
 37: em\_htg\_vrt:\*  
 38: em\_sy:\*  
 39: em\_htgo\_hum:\*  
 40: em\_htgo\_mus:\*  
 41: em\_htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a  
 score greater than or equal to the score of the result being printed,  
 and is derived by analysis of the total score distribution.

#### SUMMARIES

Result		%					
No.	Score	Query	Match	Length	DB	ID	Description
1	539	100.0	539	6	AX147744		AX147744 Sequence
2	539	100.0	539	6	AX300781		AX300781 Sequence
3	535.8	99.4	710	10	RATIGFIA		M15480 Rat insulin
4	516.6	95.8	798	10	RNIGFI2		X06108 Rat mRNA (c
5	516.6	95.8	958	10	RNIGFI1		X06107 Rat mRNA (c
6	470.2	87.2	1536	10	BC012409		BC012409 Mus muscu
7	453.8	84.2	651	10	MMIGFIBR		X04482 Mouse mRNA
8	425	78.8	487	6	AX147752		AX147752 Sequence
9	425	78.8	487	6	AX300789		AX300789 Sequence
10	421.8	78.3	1346	10	RATIGFIB		M15481 Rat insulin
11	402.6	74.7	1052	10	RATIGF1A		D00698 Rattus sp.
12	401	74.4	513	26	RNMGGJIL		M17714 Rat insulin
13	401	74.4	521	10	RNIGFIR		X06043 Rattus norv
14	387	71.8	826	10	RATIGFIAA		M17335 Rat insulin
15	361.6	67.1	696	10	MMIGFIAR		X04480 Mouse mRNA
16	356.8	66.2	523	6	AX147746		AX147746 Sequence
17	356.8	66.2	523	6	AX300783		AX300783 Sequence
18	325.2	60.3	517	6	AX147742		AX147742 Sequence
19	325.2	60.3	517	6	AX300779		AX300779 Sequence
20	299.6	55.6	671	6	AX526045		AX526045 Sequence
21	274.2	50.9	7260	6	AX375028		AX375028 Sequence
22	274.2	50.9	7260	6	AX411095		AX411095 Sequence
23	274.2	50.9	7260	9	HSIGFACI		X57025 Human IGF-I
24	272.6	50.6	666	6	A29119		A29119 H.sapiens I
25	272.6	50.6	725	9	HSIGFI		X00173 Homo sapien
26	272.6	50.6	728	9	HUMGFII		M29644 Human insul
27	271.2	50.3	616	9	HSIGF1A		X56773 H.sapiens m
28	271	50.3	620	6	I08370		I08370 Sequence 2
29	270	50.1	1076	9	HUMIGFI		M27544 Human insul
30	268.8	49.9	444	9	HSU40870		U40870 Human alter
31	267.2	49.6	432	4	AF022961		AF022961 Oryctolag
32	262	48.6	471	6	AX147754		AX147754 Sequence
33	262	48.6	471	6	AX300791		AX300791 Sequence

34	260	48.2	567	4	PIGGFIIA	M31175 Pig insulin
35	258	47.9	532	4	SSILGF1M	X17492 Porcine mRN
36	237.4	44.0	888	4	ECU85272	U85272 Equus cabal
37	231	42.9	730	9	HSIGF1B	X56774 H.sapiens m
38	231	42.9	1094	9	HUMGF1B	M11568 Human insul
39	231	42.9	1136	6	E01349	E01349 cDNA encodi
40	231	42.9	1136	6	I08009	I08009 Sequence 3
41	230.2	42.7	978	4	GOTIGFI	D11378 Goat mRNA f
42	230.2	42.7	978	6	E05279	E05279 DNA encodin
43	229.2	42.5	384	10	AF440694	AF440694 Mus muscu
44	228.6	42.4	836	10	CPIGF1	X52951 Guinea pig
45	228.6	42.4	1284	4	BTILGF1A	X15726 Bovine mRNA

# ALIGNMENTS

## RESULT 1

AX147744

LOCUS AX147744 539 bp DNA linear PAT 31-AUG-2001

DEFINITION Sequence 3 from Patent WO0136483.

ACCESSION AX147744

VERSION AX147744.1 GI:14346789

## KEYWORDS

SOURCE Rattus norvegicus (Norway rat)

ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

## REFERENCE 1

AUTHORS Goldspink, G.R. and Johnson, I.R.

TITLE Use of the insulin-like-growth factor i isoform mgf for the  
treatment of neurological disorders

JOURNAL Patent: WO 0136483-A 3 25-MAY-2001;  
University College London (GB)

## FEATURES Location/Qualifiers

source 1..539

/organism="Rattus norvegicus"

/mol\_type="genomic DNA"

/db\_xref="taxon:10116"

CDS <1..336

/note="unnamed protein product"

/codon\_start=1

/protein\_id="CAC41176.1"

/db\_xref="GI:14346790"

/db\_xref="REMTREMBL:CAC41176"

/translation="GPETLCGAELVDALQFVCGPRGFYFNKPTVYGSSIRRAPQTGIV  
DECCFRSCDLRRLEMYCVRCKPTKSARSIRAQRHTDMPKTQKSQPLSTHKRKLQRRR  
KGSTLEEHK"

BASE COUNT 161 a 136 c 139 g 103 t

## ORIGIN

Query Match 100.0%; Score 539; DB 6; Length 539;

Best Local Similarity 100.0%; Pred. No. 5.5e-157;

Matches 539; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGACCAGAGACCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTCGTGTGTGGACCA 60

```

      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1  GGACCAGAGACCCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTCGTGTGTGGACCA 60
Qy      61  AGGGGCTTTTACTTCAACAAGCCACAGTCTATGGCTCCAGCATTTCGGAGGGCACCACAG 120
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61  AGGGGCTTTTACTTCAACAAGCCACAGTCTATGGCTCCAGCATTTCGGAGGGCACCACAG 120
Qy     121  ACGGGCATTGTGGATGAGTGTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 180
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121  ACGGGCATTGTGGATGAGTGTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 180
Qy     181  TGTGTCCGCTGCAAGCCTACAAAGTCAGCTCGTTCCATCCGGGCCCAGCGCCACACTGAC 240
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181  TGTGTCCGCTGCAAGCCTACAAAGTCAGCTCGTTCCATCCGGGCCCAGCGCCACACTGAC 240
Qy     241  ATGCCCAAGACTCAGAAGTCCCAGCCCCTATCGACACACAAGAAAAGGAAGCTGCAAAGG 300
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241  ATGCCCAAGACTCAGAAGTCCCAGCCCCTATCGACACACAAGAAAAGGAAGCTGCAAAGG 300
Qy     301  AGAAGGAAAGGAAGTACACTTGAAGAACACAAGTAGAGGAAGTGCAGGAAACAAGACCTA 360
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301  AGAAGGAAAGGAAGTACACTTGAAGAACACAAGTAGAGGAAGTGCAGGAAACAAGACCTA 360
Qy     361  CAGAATGTAGGAGGAGCCTCCCGAGGAACAGAAAATGCCACGTCACCGCAAGATCCTTTG 420
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361  CAGAATGTAGGAGGAGCCTCCCGAGGAACAGAAAATGCCACGTCACCGCAAGATCCTTTG 420
Qy     421  CTGCTTGAGCAACCTGCAAAACATCGGAACACCTGCCAAATATCAATAATGAGTTCAATA 480
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421  CTGCTTGAGCAACCTGCAAAACATCGGAACACCTGCCAAATATCAATAATGAGTTCAATA 480
Qy     481  TCATTTTCAGAGATGGGCATTTCCCTCAATGAAATACACAAGTAAACATTCCCGGAATTC 539
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481  TCATTTTCAGAGATGGGCATTTCCCTCAATGAAATACACAAGTAAACATTCCCGGAATTC 539

```

# RESULT 2

AX300781

LOCUS AX300781 539 bp DNA linear PAT 30-NOV-2001

DEFINITION Sequence 3 from Patent W00185781.

ACCESSION AX300781

VERSION AX300781.1 GI:17382062

KEYWORDS

SOURCE Rattus sp.

ORGANISM Rattus sp.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

REFERENCE 1

AUTHORS Goldspink, G.D. and Terenghi, G.B.

TITLE Repair of nerve damage

JOURNAL Patent: WO 0185781-A 3 15-NOV-2001;  
University College London (GB) ; East Grinstead Medical Research  
Trust (GB)

FEATURES Location/Qualifiers

source 1..539

/organism="Rattus sp."

CDS

```
/mol_type="genomic DNA"
/db_xref="taxon:10118"
<1..336
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD13041.1"
/db_xref="GI:17382063"
/translation="GPETLCGAELVDALQFVCGPRGFYFNKPTVYGSSIIRRAPQTGIV
DECCFRSCDLRRLEMYCVRCKPTKSARSIRAQRHTDMPKTQKSQPLSTHKRKLQRRR
KGSTLEEHK"
```

BASE COUNT        161 a        136 c        139 g        103 t  
ORIGIN

Query Match                    100.0%;    Score 539;    DB 6;    Length 539;  
Best Local Similarity    100.0%;    Pred. No. 5.5e-157;  
Matches 539;    Conservative    0;    Mismatches    0;    Indels    0;    Gaps    0;

```
Qy            1 GGACCAGAGACCCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTCGTGTGTGGACCA 60
              |||
Db            1 GGACCAGAGACCCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTCGTGTGTGGACCA 60

Qy            61 AGGGGCTTTTACTTCAACAAGCCACAGTCTATGGCTCCAGCATTTCGGAGGGCACCACAG 120
              |||
Db            61 AGGGGCTTTTACTTCAACAAGCCACAGTCTATGGCTCCAGCATTTCGGAGGGCACCACAG 120

Qy            121 ACGGGCATTGTGGATGAGTGTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 180
              |||
Db            121 ACGGGCATTGTGGATGAGTGTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 180

Qy            181 TGTGTCCGCTGCAAGCCTACAAAGTCAGCTCGTTCCATCCGGGGCCAGCGCCCACTGAC 240
              |||
Db            181 TGTGTCCGCTGCAAGCCTACAAAGTCAGCTCGTTCCATCCGGGGCCAGCGCCCACTGAC 240

Qy            241 ATGCCCAAGACTCAGAAGTCCCAGCCCCCTATCGACACACAAGAAAAGGAAGCTGCAAAGG 300
              |||
Db            241 ATGCCCAAGACTCAGAAGTCCCAGCCCCCTATCGACACACAAGAAAAGGAAGCTGCAAAGG 300

Qy            301 AGAAGGAAAGGAAGTACACTTGAAGAACACAAGTAGAGGAAGTGCAGGAAACAAGACCTA 360
              |||
Db            301 AGAAGGAAAGGAAGTACACTTGAAGAACACAAGTAGAGGAAGTGCAGGAAACAAGACCTA 360

Qy            361 CAGAATGTAGGAGGAGCCTCCCGAGGAACAGAAAATGCCACGTACCGCAAGATCCTTTG 420
              |||
Db            361 CAGAATGTAGGAGGAGCCTCCCGAGGAACAGAAAATGCCACGTACCGCAAGATCCTTTG 420

Qy            421 CTGCTTGAGCAACCTGCAAAACATCGGAACACCTGCCAAATATCAATAATGAGTTCAATA 480
              |||
Db            421 CTGCTTGAGCAACCTGCAAAACATCGGAACACCTGCCAAATATCAATAATGAGTTCAATA 480

Qy            481 TCATTTTCAGAGATGGGCATTTCCCTCAATGAAATACACAAGTAAACATTCCCGGAATTTC 539
              |||
Db            481 TCATTTTCAGAGATGGGCATTTCCCTCAATGAAATACACAAGTAAACATTCCCGGAATTTC 539
```

RESULT 3

RATIGFIA

LOCUS

RATIGFIA

710 bp

mRNA

linear

ROD 27-APR-1993



Db 172 GGACCAGAGACCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTTCGTGTGTGGACCA 231

Qy 61 AGGGGCTTTTACTTCAACAAGCCCACAGTCTATGGCTCCAGCATTTCGGAGGGCACCACAG 120  
 |||||

Db 232 AGGGGCTTTTACTTCAACAAGCCCACAGGCTATGGCTCCAGCATTTCGGAGGGCACCACAG 291

Qy 121 ACGGGCATTGTGGATGAGTGTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 180  
 |||||

Db 292 ACGGGCATTGTGGATGAGTGTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 351

Qy 181 TGTGTCCGCTGCAAGCCTACAAAGTCAGCTCGTTCCATCCGGGGCCAGCGCCCACTGAC 240  
 |||||

Db 352 TGTGTCCGCTGCAAGCCTACAAAGTCAGCTCGTTCCATCCGGGGCCAGCGCCCACTGAC 411

Qy 241 ATGCCCCAAGACTCAGAAGTCCCAGCCCCTATCGACACACAAGAAAAGGAAGCTGCAAAGG 300  
 |||||

Db 412 ATGCCCCAAGACTCAGAAGTCCCAGCCCCTATCGACACACAAGAAAAGGAAGCTGCAAAGG 471

Qy 301 AGAAGGAAAGGAAGTACACTTGAAGAACACAAGTAGAGGAAGTGCAGGAAACAAGACCTA 360  
 |||||

Db 472 AGAAGGAAAGGAAGTACACTTGAAGAACACAAGTAGAGGAAGTGCAGGAAACAAGACCTA 531

Qy 361 CAGAATGTAGGAGGAGCCTCCCGAGGAACAGAAAATGCCACGTCACCGCAAGATCCTTTG 420  
 |||||

Db 532 CAGAATGTAGGAGGAGCCTCCCGAGGAACAGAAAATGCCACGTCACCGCAAGATCCTTTG 591

Qy 421 CTGCTTGAGCAACCTGCAAAACATCGGAACACCTGCCAAATATCAATAATGAGTTCAATA 480  
 |||||

Db 592 CTGCTTGAGCAACCTGCAAAACATCGGAACACCTGCCAAATATCAATAATGAGTTCAATA 651

Qy 481 TCATTTTCAGAGATGGGCATTTCCCTCAATGAAATACACAAGTAAACATTCCCGGAATTC 539  
 |||||

Db 652 CCATTTTCAGAGATGGGCATTTCCCTCAATGAAATACACAAGTAAACATTCCCGGAATTC 710

#### RESULT 4

##### RNIGFI2

LOCUS RNIGFI2 798 bp mRNA linear ROD 12-SEP-1993

DEFINITION Rat mRNA (clone IGF1AB2) for insulin-like growth factor I.

ACCESSION X06108 M32339 Y00429

VERSION X06108.1 GI:56426

KEYWORDS insulin-like growth factor I.

SOURCE Rattus norvegicus (Norway rat)

ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.

##### REFERENCE 1

AUTHORS Shimatsu, A. and Rotwein, P.

TITLE Sequence of Two Rat Insulin-like Growth Factor I mRNAs Differing  
 Within the 5' Untranslated Region

JOURNAL Nucleic Acids Res. 15 (1987) In press

##### REFERENCE 2 (bases 1 to 798)

AUTHORS Rotwein, P.

TITLE Direct Submission

JOURNAL Submitted (21-OCT-1987) Rotwein P., Washington University, School  
 of Medicine, 660 South Euclid Avenue, Box 8127, St. Louis, MO

```

63110, USA
COMMENT      Another IGF-I mRNA of rat liver differing in the 5' UT-region is
              described in <X06107>.
FEATURES             Location/Qualifiers
     source             1. .798
                       /organism="Rattus norvegicus"
                       /mol_type="mRNA"
                       /strain="Sprague-Dawley"
                       /db_xref="taxon:10116"
                       /tissue_type="liver"
     misc_feature       1. .72
                       /note="5' UT-region"
     CDS                73. .552
                       /note="IGF-I (AA 1-159) "
                       /codon_start=1
                       /protein_id="CAA29481.1"
                       /db_xref="GI:56427"
                       /translation="MGKISSLPTQLFKICLDFLKIKIHIMSSSHLFYLALCLLTFTS
SATAGPETLCGAELVDALQFVCGPRGFYFNKPTGYGSSIRRAPQTGIVDECCFRSCDL
RRLEMYCAPLKPTKSARSIRAQRHTDMPKTQKSQPLSTHKKRKLQRRRKGSTLEEHK"
     misc_feature       553. .798
                       /note="3' UT-region"
BASE COUNT      238 a      196 c      183 g      181 t
ORIGIN

```

Qy	1	GGACCAGAGACCCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTCGTGTGTGGACCA	60
Db	217	GGACCAGAGACCCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTCGTGTGTGGACCA	276
Qy	61	AGGGGCTTTTACTTCAACAAGCCCACAGTCTATGGCTCCAGCATTCCGAGGGCACCAACAG	120
Db	277	AGGGGCTTTTACTTCAACAAGCCCACAGGCTATGGCTCCAGCATTCCGAGGGCACCAACAG	336
Qy	121	ACGGGCATTGTGGATGAGTGTTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC	180
Db	337	ACGGGCATTGTGGATGAGTGTTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC	396
Qy	181	TGTGTCCGCTGCAAGCCTACAAAGTCAGCTCGTTCCATCCGGGCCACGCCCACACTGAC	240
Db	397	TGTGCTCCGCTGAAGCCTACAAAGTCAGCTCGTTCCATCCGGGCCACGCCCACACTGAC	456
Qy	241	ATGCCCAAGACTCAGAAGTCCAGCCCCCTATCGACACACAAGAAAAGGAAGCTGCAAAGG	300
Db	457	ATGCCCAAGACTCAGAAGTCCAGCCCCCTATCGACACACAAGAAAAGGAAGCTGCAAAGG	516
Qy	301	AGAAGGAAAGGAAGTACACTTGAAGAACACAAGTAGAGGAAGTGCAGGAAACAAGACCTA	360
Db	517	AGAAGGAAAGGAAGTACACTTGAAGAACACAAGTAGAGGAAGTGCAGGAAACAAGACCTA	576
Qy	361	CAGAATGTAGGAGGAGCCTCCCGAGGAACAGAAAAATGCCACGTACCGCAAGATCCTTTG	420
Db	577	CAGAATGTAGGAGGAGCCTCCCGAGGAACAGAAAAATGCCACGTACCGCAAGATCCTTTG	636



Qy 421 CTGCTTGAGCAACCTGCAAAACATCGGAACACCTGCCAAATATCAATAATGAGTTCAATA 480  
 |||  
 Db 637 CTGCTTGAGCAACCTGCAAAACATCGGAACACCTGCCAAATATCAATAATGAGTTCAATA 696

Qy 481 TCATTTTCAGAGATGGGCATTTCCCTCAATGAAATACACAAGTAAACATTCC 531  
 |||  
 Db 697 CCATTTTCAGAGATGGGCATTTCCCTCAATGAAATACACAAGTAAACATTCC 747

# RESULT 5

RNIGFI1

LOCUS RNIGFI1 958 bp mRNA linear ROD 12-SEP-1993

DEFINITION Rat mRNA (clone IGF1AB1) for insulin-like growth factor I.

ACCESSION X06107 M32260 Y00429

VERSION X06107.1 GI:56424

KEYWORDS insulin-like growth factor I.

SOURCE Rattus norvegicus (Norway rat)

ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.

## REFERENCE

1

AUTHORS Shimatsu,A. and Rotwein,P.

TITLE Sequence of two rat insulin-like growth factor I mRNAs differing  
 within the 5' untranslated region

JOURNAL Nucleic Acids Res. 15 (17), 7196 (1987)

MEDLINE 88015572

PUBMED 3658684

REFERENCE 2 (bases 1 to 958)

AUTHORS Rotwein,P.

TITLE Direct Submission

JOURNAL Submitted (21-OCT-1987) Rotwein P., Washington University, School  
 of Medicine, 660 South Euclid Avenue, Box 8127, St. Louis, MO  
 63110, USA

COMMENT Another IGF-I mRNA of rat liver differing in the 5' UT-region is  
 described in <X06108>.

FEATURES Location/Qualifiers

source

1. .958

/organism="Rattus norvegicus"

/mol\_type="mRNA"

/strain="Srpague-Dawley"

/db\_xref="taxon:10116"

/tissue\_type="liver"

misc\_feature

1. .258

/note="5' UT-region"

CDS

259. .738

/note="IGF-I (AA 1-159)"

/codon\_start=1

/protein\_id="CAA29480.1"

/db\_xref="GI:56425"

/db\_xref="SWISS-PROT:P08024"

/translation="MGKISSLPQLFKICLCDFLKIKIHIMSSSHLFYLALCLLTFTS  
 SATAGPETLCGAELVDALQFVCGPRGFYFNKPTGYGSSIRRAPQTGIVDECCFRSCDL  
 RRLEMYCAPLKPTKSARSIRAQRHTDMPKTQKSQPLSTHKRKLQRRRKGSTLEEHK"

misc\_feature

736. .958

/note="3' UT-region"

BASE COUNT 260 a 246 c 209 g 243 t

# ORIGIN

Query Match 95.8%; Score 516.6; DB 10; Length 958;  
 Best Local Similarity 98.3%; Pred. No. 6.1e-150;  
 Matches 522; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

```

Qy      1 GGACCAGAGACCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTCGTGTGTGGACCA 60
      |||
Db      403 GGACCAGAGACCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTCGTGTGTGGACCA 462

Qy      61 AGGGGCTTTTACTTCAACAAGCCACAGTCTATGGCTCCAGCATTTCGAGGGCACCACAG 120
      |||
Db      463 AGGGGCTTTTACTTCAACAAGCCACAGGCTATGGCTCCAGCATTTCGAGGGCACCACAG 522

Qy      121 ACGGGCATTGTGGATGAGTGTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 180
      |||
Db      523 ACGGGCATTGTGGATGAGTGTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 582

Qy      181 TGTGTCCGCTGCAAGCCTACAAAGTCAGCTCGTTCCATCCGGGGCCAGCGCCACACTGAC 240
      |||
Db      583 TGTGTCCGCTGCAAGCCTACAAAGTCAGCTCGTTCCATCCGGGGCCAGCGCCACACTGAC 642

Qy      241 ATGCCCAAGACTCAGAAGTCCCAGCCCCTATCGACACACAAGAAAAGGAAGCTGCAAAGG 300
      |||
Db      643 ATGCCCAAGACTCAGAAGTCCCAGCCCCTATCGACACACAAGAAAAGGAAGCTGCAAAGG 702

Qy      301 AGAAGGAAAGGAAGTACACTTGAAGAACACAAGTAGAGGAAGTGCAGGAAACAAGACCTA 360
      |||
Db      703 AGAAGGAAAGGAAGTACACTTGAAGAACACAAGTAGAGGAAGTGCAGGAAACAAGACCTA 762

Qy      361 CAGAATGTAGGAGGAGCCTCCCGAGGAACAGAAAATGCCACGTACCGCAAGATCCTTTG 420
      |||
Db      763 CAGAATGTAGGAGGAGCCTCCCGAGGAACAGAAAATGCCACGTACCGCAAGATCCTTTG 822

Qy      421 CTGCTTGAGCAACCTGCAAAACATCGGAACACCTGCCAAATATCAATAATGAGTTCAATA 480
      |||
Db      823 CTGCTTGAGCAACCTGCAAAACATCGGAACACCTGCCAAATATCAATAATGAGTTCAATA 882

Qy      481 TCATTTTCAGAGATGGGCATTTCCTCAATGAAATACACAAGTAAACATTCC 531
      |||
Db      883 CCATTTTCAGAGATGGGCATTTCCTCAATGAAATACACAAGTAAACATTCC 933
  
```

## RESULT 6

BC012409

LOCUS BC012409 1536 bp mRNA linear ROD 16-APR-2003

DEFINITION Mus musculus insulin-like growth factor 1, mRNA (cDNA clone MGC:18617 IMAGE:4194295), complete cds.

ACCESSION BC012409

VERSION BC012409.1 GI:15214568

KEYWORDS MGC.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 1536)

AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,

Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smailus, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.

TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

MEDLINE 22388257

PUBMED 12477932

REFERENCE 2 (bases 1 to 1536)

AUTHORS Strausberg, R.

TITLE Direct Submission

JOURNAL Submitted (15-AUG-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>

COMMENT Contact: MGC help desk  
Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
Tissue Procurement: Jeffrey E. Green, M.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center  
Center code: BCM-HGSC  
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>  
Contact: [amg@bcm.tmc.edu](mailto:amg@bcm.tmc.edu)  
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H., Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati, A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAK Plate: 24 Row: k Column: 11  
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 6754307.

FEATURES

source Location/Qualifiers

1. .1536

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="FVB/N"

/db\_xref="taxon:10090"

/clone="MGC:18617 IMAGE:4194295"

/tissue\_type="Liver, normal. 5 month old male mouse."

```

/clone_lib="NCI_CGAP_Li9"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
gene 1. 1536
      /gene="Igf1"
      /db_xref="LocusID:16000"
      /db_xref="MGI:96432"
CDS   28. 429
      /codon_start=1
      /product="insulin-like growth factor 1"
      /protein_id="AAH12409.1"
      /db_xref="GI:15214569"
      /db_xref="LocusID:16000"
      /translation="MSSSHLFYLALCLLTFTSSTTAGPETLCGAELVDALQFVCGPRG
FYFNKPTGYGSSIRRAPQTGIVDECCFRSCDLRRLEMYCAPLKPTKAARSIRAQRHTD
MPKTQKSPSLSTNKKTKLQRRRKGSTFEEHK"
BASE COUNT      485 a      324 c      303 g      424 t
ORIGIN

```

```

Query Match          87.2%;  Score 470.2;  DB 10;  Length 1536;
Best Local Similarity 92.8%;  Pred. No. 2.1e-135;
Matches 493;  Conservative 0;  Mismatches 38;  Indels 0;  Gaps 0;

```

```

Qy      1 GGACCAGAGACCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTCGTGTGTGGACCA 60
      |||
Db      94 GGACCAGAGACCCTTTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGACCG 153
      |||

Qy      61 AGGGGCTTTTACTTCAACAAGCCACAGTCTATGGCTCCAGCATTCCGAGGGCACCACAG 120
      |||
Db     154 AGGGGCTTTTACTTCAACAAGCCACAGGCTATGGCTCCAGCATTCCGAGGGCACCTCAG 213
      |||

Qy     121 ACGGGCATTGTGGATGAGTGTTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 180
      ||
Db     214 ACAGGCATTGTGGATGAGTGTTGCTTCCGGAGCTGTGATCTGAGGAGACTGGAGATGTAC 273
      |||

Qy     181 TGTGTCCGCTGCAAGCCTACAAAGTCAGCTCGTTCCATCCGGGCCCAGCGCCCACTGAC 240
      ||||
Db     274 TGTGCCCCACTGAAGCCTACAAAGCAGCCCGCTCTATCCGTGCCAGCGCCCACTGAC 333
      ||||

Qy     241 ATGCCCAAGACTCAGAAGTCCCAGCCCCTATCGACACACAAGAAAAGGAAGCTGCAAAGG 300
      |||
Db     334 ATGCCCAAGACTCAGAAGTCCCCGTCCCTATCGACAAACAAGAAAACGAAGCTGCAAAGG 393
      |||

Qy     301 AGAAGGAAAGGAAGTACACTTGAAGAACACAAGTAGAGGAAGTGCAGGAAACAAGACCTA 360
      |||
Db     394 AGAAGGAAAGGAAGTACATTTGAAGAACACAAGTAGAGGAAGTGCAGGAAACAAGACCTA 453
      |||

Qy     361 CAGAATGTAGGAGGAGCCTCCCGAGGAACAGAAAATGCCACGTCACCGCAAGATCCTTTG 420
      |||
Db     454 CAGAATGTAGGAGGAGCCTCCCGAGGAGCAGAAAATGCCACATCACCGCAGGATCCTTTG 513
      |||

Qy     421 CTGCTTGAGCAACCTGCAAAACATCGGAACACCTGCCAAATATCAATAATGAGTTCAATA 480
      |||
Db     514 CTGCTTGAGCAACCTGCAAAACATCGAAACACCTACCAAATAACAATAATAAGTCCAATA 573
      |||

Qy     481 TCATTTTCAGAGATGGGCATTTCCCTCAATGAAATACACAAGTAAACATTCC 531
      |||

```

Db

574 ACATTACAAAGATGGGCATTTCCCCCAATGAAATATACAAGTAAACATTCC 624

## RESULT 7

## MMIGFIBR

LOCUS MMIGFIBR 651 bp mRNA linear ROD 21-MAR-1995

DEFINITION Mouse mRNA for preproinsulin-like growth factor IB.

ACCESSION X04482

VERSION X04482.1 GI:51806

KEYWORDS growth factor; insulin-like growth factor IB; preproinsulin-like growth factor IB; signal peptide.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 651)

AUTHORS Bell,G.I., Stempien,M.M., Fong,N.M. and Rall,L.B.

TITLE Sequences of liver cDNAs encoding two different mouse insulin-like growth factor I precursors

JOURNAL Nucleic Acids Res. 14 (20), 7873-7882 (1986)

MEDLINE 87040760

PUBMED 3774549

COMMENT The sequence is identical to the preproIGF-IA sequence (X04480) except for the presence of a 52 bp insertion following codon 86 (position 397 to 448), caused by alternative RNA splicing. The B domain of IGF comprises residues 1-29 (position 139-225), the C domain residues 30-41 (position 226-261), the A domain residues 42-62 (position 262-324) and the D domain residues 63-70 (position 325-348).

## FEATURES

source

Location/Qualifiers

1. .651

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="Balb/c"

/db\_xref="taxon:10090"

/clone="migfl-4"

/tissue\_type="liver"

misc\_feature

43. .45

/note="pot. translation start"

CDS

73. .474

/note="prepro IGF-IB (aa -22 to 111)"

/codon\_start=1

/protein\_id="CAA28170.1"

/db\_xref="GI:51807"

/db\_xref="MGI:96432"

/db\_xref="SWISS-PROT:P05018"

/translation="MSSSHLFYLA LCLLTFTSSTTAGPETLCGAELVDALQFVCGPRG  
FYFNKPTGYGSSIRRAPQTGIVDECCFRSCDLRRLEMYCAPLKPTKAARSIRAQRHTD  
MPKTQKSPSLSTNKKTKLQRRRKGSTFEEHK"

sig\_peptide

73. .138

/note="signal peptide (aa -22 to -1)"

mat\_peptide

139. .348

/product="mature IGF-IB (aa 1-70)"

misc\_feature

349. .471

/note="COOH-terminal peptide (E domain) (aa 71 to 111)"

polyA\_site

651

/note="polyA site"

BASE COUNT        193 a       185 c       149 g       124 t  
ORIGIN

Query Match                    84.2%;    Score 453.8;    DB 10;    Length 651;  
Best Local Similarity    92.8%;    Pred. No. 2.4e-130;  
Matches 476;    Conservative    0;    Mismatches    37;    Indels       0;    Gaps       0;

```
Qy      1 GGACCAGAGACCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTCGTGTGTGGACCA 60
      |||
Db     139 GGACCAGAGACCCTTTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGACCG 198

Qy     61 AGGGGCTTTTACTTCAACAAGCCCACAGTCTATGGCTCCAGCATTTCGGAGGGCACCACAG 120
      |||
Db     199 AGGGGCTTTTACTTCAACAAGCCCACAGGCTATGGCTCCAGCATTTCGGAGGGCACCTCAG 258

Qy    121 ACGGGCATTGTGGATGAGTGTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 180
      ||
Db    259 ACAGGCATTGTGGATGAGTGTGCTTCCGGAGCTGTGATCTGAGGAGACTGGAGATGTAC 318

Qy    181 TGTGTCCGCTGCAAGCCTACAAAGTCAGCTCGTTCCATCCGGGCCCAGCGCCCACTGAC 240
      ||||
Db    319 TGTGCCCCACTGAAGCCTACAAAAGCAGCCCGCTCTATCCGTGCCAGCGCCCACTGAC 378

Qy    241 ATGCCCAAGACTCAGAAGTCCCAGCCCCTATCGACACACAAGAAAAGGAAGCTGCAAAGG 300
      |||
Db    379 ATGCCCAAGACTCAGAAGTCCCGTCCCTATCGACAAACAAGAAAACGAAGCTGCAAAGG 438

Qy    301 AGAAGGAAAGGAAGTACACTTGAAGAACACAAGTAGAGGAAGTGCAGGAAACAAGACCTA 360
      |||
Db    439 AGAAGGAAAGGAAGTACATTTGAAGAACACAAGTAGAGGAAGTGCAGGAAACAAGACCTA 498

Qy    361 CAGAATGTAGGAGGAGCCTCCCGAGGAACAGAAAATGCCACGTCAACGCAAGATCCTTTG 420
      |||
Db    499 CAGAATGTAGGAGGAGCCTCCCGAGGAGCAGAAAATGCCACATCAACGCAAGATCCTTTG 558

Qy    421 CTGCTTGAGCAACCTGCAAAACATCGGAACACCTGCCAAATATCAATAATGAGTTCAATA 480
      |||
Db    559 CTGCTTGAGCAACCTGCAAAACATCGAAACACCTACCAAATAACAATAATAAGTCCAATA 618

Qy    481 TCATTTTCAGAGATGGGCATTTCCCTCAATGAAA 513
      |||
Db    619 ACATTACAAAGATGGGCATTTCCCCCAATGAAA 651
```

RESULT 8  
AX147752

LOCUS            AX147752                    487 bp       DNA       linear       PAT 31-AUG-2001  
DEFINITION      Sequence 11 from Patent WO0136483.  
ACCESSION       AX147752  
VERSION          AX147752.1    GI:14346797  
KEYWORDS  
SOURCE           Rattus norvegicus (Norway rat)  
         ORGANISM    Rattus norvegicus  
                     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
                     Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
                     Rattus.  
REFERENCE        1

AUTHORS Goldspink,G.R. and Johnson,I.R.  
 TITLE Use of the insulin-like-growth factor i isoform mgf for the  
 treatment of neurological disorders  
 JOURNAL Patent: WO 0136483-A 11 25-MAY-2001;  
 University College London (GB)  
 FEATURES Location/Qualifiers  
     source 1. .487  
             /organism="Rattus norvegicus"  
             /mol\_type="genomic DNA"  
             /db\_xref="taxon:10116"  
     CDS <1. .318  
             /note="unnamed protein product"  
             /codon\_start=1  
             /protein\_id="CAC41179.1"  
             /db\_xref="GI:14346798"  
             /db\_xref="REMTREMBL:CAC41179"  
             /translation="GPETLCGAELVDALQFVCGPRGFYFNKPTVYGSSIRRAPQTGIV  
             DECCFRSCDLRRLEMYCVRCKPTKSARSIRAQRHTDMPKTQKEVHLKNTSRGSAGNKT  
             YRM"  
 BASE COUNT 139 a 123 c 126 g 99 t  
 ORIGIN

Query Match 78.8%; Score 425; DB 6; Length 487;  
 Best Local Similarity 90.4%; Pred. No. 2.3e-121;  
 Matches 487; Conservative 0; Mismatches 0; Indels 52; Gaps 1;

```

Qy      1 GGACCAGAGACCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTCGTGTGTGGACCA 60
      |||
Db      1 GGACCAGAGACCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTCGTGTGTGGACCA 60

Qy     61 AGGGGCTTTTACTTCAACAAGCCACAGTCTATGGCTCCAGCATTTCGGAGGGCACCACAG 120
      |||
Db     61 AGGGGCTTTTACTTCAACAAGCCACAGTCTATGGCTCCAGCATTTCGGAGGGCACCACAG 120

Qy    121 ACGGGCATTGTGGATGAGTGTGTGCTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 180
      |||
Db    121 ACGGGCATTGTGGATGAGTGTGTGCTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 180

Qy    181 TGTGTCCGCTGCAAGCCTACAAAGTCAGCTCGTTCCATCCGGGCCCAGCGCCCACTGAC 240
      |||
Db    181 TGTGTCCGCTGCAAGCCTACAAAGTCAGCTCGTTCCATCCGGGCCCAGCGCCCACTGAC 240

Qy    241 ATGCCCAAGACTCAGAAGTCCCAGCCCCTATCGACACACAAGAAAAGGAAGCTGCAAAGG 300
      |||
Db    241 ATGCCCAAGACTCAG----- 255

Qy    301 AGAAGGAAAGGAAGTACACTTGAAGAACACAAGTAGAGGAAGTGCAGGAAACAAGACCTA 360
      |||
Db    256 -----AAGGAAGTACACTTGAAGAACACAAGTAGAGGAAGTGCAGGAAACAAGACCTA 308

Qy    361 CAGAATGTAGGAGGAGCCTCCCGAGGAACAGAAAATGCCACGTACCGCAAGATCCTTTG 420
      |||
Db    309 CAGAATGTAGGAGGAGCCTCCCGAGGAACAGAAAATGCCACGTACCGCAAGATCCTTTG 368

Qy    421 CTGCTTGAGCAACCTGCAAAACATCGGAACACCTGCCAAATATCAATAATGAGTTCAATA 480
      |||
Db    369 CTGCTTGAGCAACCTGCAAAACATCGGAACACCTGCCAAATATCAATAATGAGTTCAATA 428

```

Qy 481 TCATTTTCAGAGATGGGCATTTCCTCAATGAAATACACAAGTAAACATTCCCAGGAATTC 539  
 |||  
 Db 429 TCATTTTCAGAGATGGGCATTTCCTCAATGAAATACACAAGTAAACATTCCCAGGAATTC 487

RESULT 9

AX300789

LOCUS AX300789 487 bp DNA linear PAT 30-NOV-2001

DEFINITION Sequence 11 from Patent WO0185781.

ACCESSION AX300789

VERSION AX300789.1 GI:17382070

KEYWORDS

SOURCE Rattus sp.

ORGANISM Rattus sp.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.

REFERENCE 1

AUTHORS Goldspink, G.D. and Terenghi, G.B.

TITLE Repair of nerve damage

JOURNAL Patent: WO 0185781-A 11 15-NOV-2001;  
 University College London (GB) ; East Grinstead Medical Research  
 Trust (GB)

FEATURES Location/Qualifiers

source

1..487

/organism="Rattus sp."

/mol\_type="genomic DNA"

/db\_xref="taxon:10118"

CDS

<1..318

/note="unnamed protein product"

/codon\_start=1

/protein\_id="CAD13044.1"

/db\_xref="GI:17382071"

/translation="GPETLCGAELVDALQFVCGPRGFYFNKPTVYGSSIRRAPQTGIV  
 DECCFRSCDLRRLMEYCVRCRPTKSARSIRAQRHTDMPKTQKEVHLKNTSRGSAGNKT  
 YRM"

BASE COUNT 139 a 123 c 126 g 99 t

ORIGIN

Query Match 78.8%; Score 425; DB 6; Length 487;

Best Local Similarity 90.4%; Pred. No. 2.3e-121;

Matches 487; Conservative 0; Mismatches 0; Indels 52; Gaps 1;

Qy 1 GGACCAGAGACCCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTTCGTGTGTGGACCA 60  
 |||  
 Db 1 GGACCAGAGACCCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTTCGTGTGTGGACCA 60

Qy 61 AGGGGCTTTTACTTCAACAAGCCACAGTCTATGGCTCCAGCATTTCGGAGGGCACCACAG 120  
 |||  
 Db 61 AGGGGCTTTTACTTCAACAAGCCACAGTCTATGGCTCCAGCATTTCGGAGGGCACCACAG 120

Qy 121 ACGGGCATTGTGGATGAGTGTTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 180  
 |||  
 Db 121 ACGGGCATTGTGGATGAGTGTTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 180

Qy 181 TGTGTCCGCTGCAAGCCTACAAAGTCAGCTCGTTCCATCCGGGCCAGCGCCACACTGAC 240



```

      |||
Db      181 TGTGTCCGCTGCAAGCCTACAAAGTCAGCTCGTTCCATCCGGGCCCAGCGCCACACTGAC 240
Qy      241 ATGCCCAAGACTCAGAAGTCCCAGCCCCTATCGACACACAAGAAAAGGAAGCTGCAAAGG 300
      |||
Db      241 ATGCCCAAGACTCAG----- 255
Qy      301 AGAAGGAAAGGAAGTACACTTGAAGAACACAAGTAGAGGAAGTGCAGGAAACAAGACCTA 360
      |||
Db      256 -----AAGGAAGTACACTTGAAGAACACAAGTAGAGGAAGTGCAGGAAACAAGACCTA 308
Qy      361 CAGAATGTAGGAGGAGCCTCCCGAGGAACAGAAAATGCCACGTCACCGCAAGATCCTTTG 420
      |||
Db      309 CAGAATGTAGGAGGAGCCTCCCGAGGAACAGAAAATGCCACGTCACCGCAAGATCCTTTG 368
Qy      421 CTGCTTGAGCAACCTGCAAAACATCGGAACACCTGCCAAATATCAATAATGAGTTCAATA 480
      |||
Db      369 CTGCTTGAGCAACCTGCAAAACATCGGAACACCTGCCAAATATCAATAATGAGTTCAATA 428
Qy      481 TCATTTTCAGAGATGGGCATTTCCTCAATGAAATACACAAGTAAACATTCCCGGAATTC 539
      |||
Db      429 TCATTTTCAGAGATGGGCATTTCCTCAATGAAATACACAAGTAAACATTCCCGGAATTC 487

```

# RESULT 10

## RATIGFIB

LOCUS RATIGFIB 1346 bp mRNA linear ROD 27-APR-1993

DEFINITION Rat insulin-like growth factor I (IGF-I) mRNA, complete cds.

ACCESSION M15481

VERSION M15481.1 GI:204753

KEYWORDS growth factor; insulin-like growth factor.

SOURCE Rattus norvegicus (Norway rat)

ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

REFERENCE 1 (bases 1 to 1346)

AUTHORS Roberts,C.T. Jr., Lasky,S.R., Lowe,W.L. Jr., Seaman,W.T. and  
LeRoith,D.

TITLE Molecular cloning of rat insulin-like growth factor I complementary  
deoxyribonucleic acids: differential messenger ribonucleic acid  
processing and regulation by growth hormone in extrahepatic tissues

JOURNAL Mol. Endocrinol. 1 (3), 243-248 (1987)

MEDLINE 88288198

PUBMED 3453891

COMMENT Original source text: Rat (Sprague-Dawley) adult liver, cDNA to  
mRNA, clone pRIGF-1-25.

Draft entry and computer-readable copy of sequence in [Mol.  
Endocrinol. (1987) In press] kindly  
provided by S.R.Lasky, 16-MAR-1987.

FEATURES Location/Qualifiers

source 1..1346  
/organism="Rattus norvegicus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10116"

CDS 794..1177  
/note="insulin-like growth factor precursor"

```

/codon_start=1
/protein_id="AAA41387.1"
/db_xref="GI:204754"
/translation="MSSSHLFYIALCLLTFTSSATAGPETLCGAELVDALQFVCGPRG
FYFNKPTGYGSSIRRAPQTGIVDECCFRSCDLRRLEMYCVRCKPTKSARSIRAQRHTD
MPKTQKEVHLKNTSRGSAGNKTYRM"
sig_peptide      794. .859
                  /note="insulin-like growth factor signal peptide"
mat_peptide      860. .946
                  /product="insulin-like growth factor B peptide"
mat_peptide      947. .982
                  /product="insulin-like growth factor C peptide"
mat_peptide      983. .1045
                  /product="insulin-like growth factor A peptide"
mat_peptide      1046. .1069
                  /product="insulin-like growth factor D peptide"
mat_peptide      1070. .1174
                  /product="insulin-like growth factor E peptide"
BASE COUNT      348 a    344 c    344 g    310 t
ORIGIN          1 bp upstream of EcoRI site.

```

```

Query Match      78.3%; Score 421.8; DB 10; Length 1346;
Best Local Similarity 90.0%; Pred. No. 2.8e-120;
Matches 485; Conservative 0; Mismatches 2; Indels 52; Gaps 1;

```

```

Qy      1 GGACCAGAGACCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTCGTGTGTGGACCA 60
      |||
Db      860 GGACCAGAGACCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTCGTGTGTGGACCA 919
      |||
Qy      61 AGGGGCTTTTACTTCAACAAGCCACAGTCTATGGCTCCAGCATTTCGGAGGGCACCACAG 120
      |||
Db      920 AGGGGCTTTTACTTCAACAAGCCACAGGCTATGGCTCCAGCATTTCGGAGGGCACCACAG 979
      |||
Qy      121 ACGGGCATTGTGGATGAGTGTGTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 180
      |||
Db      980 ACGGGCATTGTGGATGAGTGTGTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 1039
      |||
Qy      181 TGTGTCCGCTGCAAGCCTACAAAGTCAGCTCGTTCCATCCGGGCCCAGCGCCCACTGAC 240
      |||
Db      1040 TGTGTCCGCTGCAAGCCTACAAAGTCAGCTCGTTCCATCCGGGCCCAGCGCCCACTGAC 1099
      |||
Qy      241 ATGCCCAAGACTCAGAAGTCCCAGCCCCATCGACACACAAGAAAAGGAAGCTGCAAAGG 300
      |||
Db      1100 ATGCCCAAGACTCAG----- 1114
      |||
Qy      301 AGAAGGAAAGGAAGTACACTTGAAGAACACAAGTAGAGGAAGTGAGGAAACAAGACCTA 360
      |||
Db      1115 -----AAGGAAGTACACTTGAAGAACACAAGTAGAGGAAGTGAGGAAACAAGACCTA 1167
      |||
Qy      361 CAGAATGTAGGAGGAGCCTCCCGAGGAACAGAAAATGCCACGTCACCGCAAGATCCTTTG 420
      |||
Db      1168 CAGAATGTAGGAGGAGCCTCCCGAGGAACAGAAAATGCCACGTCACCGCAAGATCCTTTG 1227
      |||
Qy      421 CTGCTTGAGCAACCTGCAAAACATCGGAACACCTGCCAAATATCAATAATGAGTTCAATA 480
      |||
Db      1228 CTGCTTGAGCAACCTGCAAAACATCGGAACACCTGCCAAATATCAATAATGAGTTCAATA 1287
      |||

```

Qy 481 TCATTTTCAGAGATGGGCATTTCCCTCAATGAAATACACAAGTAAACATTCCCAGGAATTC 539  
 |||  
 Db 1288 CCATTTTCAGAGATGGGCATTTCCCTCAATGAAATACACAAGTAAACATTCCCAGGAATTC 1346

# RESULT 11

## RATIGF1A

LOCUS RATIGF1A 1052 bp mRNA linear ROD 29-MAY-2002

DEFINITION Rattus sp. mRNA for insulin-like growth factor I precursor,  
 complete cds.

ACCESSION D00698

VERSION D00698.1 GI:220780

KEYWORDS insulin-like growth factor I.

SOURCE Rattus sp.

ORGANISM Rattus sp.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.

REFERENCE 1 (bases 28 to 1052)

AUTHORS Shimatsu,A. and Rotwein,P.

TITLE Mosaic evolution of the insulin-like growth factors. Organization,  
 sequence, and expression of the rat insulin-like growth factor I  
 gene

JOURNAL J. Biol. Chem. 262 (16), 7894-7900 (1987)

MEDLINE 87222423

PUBMED 3034909

REFERENCE 2 (bases 1 to 1052)

AUTHORS Kato,H., Okoshi,A., Miura,Y. and Noguchi,T.

TITLE A new cDNA clone relating to larger molecular species of rat  
 insulin-like growth factor-I mRNA

JOURNAL Agric. Biol. Chem. 54 (6), 1599-1601 (1990)

MEDLINE 91103966

PUBMED 1368571

COMMENT The difference in the size of IGF-I mRNA has been suggested to be  
 primarily due to differences in the 3'-untranslated region in [1].  
 The cDNA clone to IGF-I mRNA reported in [2] gives direct evidence  
 for this hypothesis.

FEATURES Location/Qualifiers

source

1..1052  
 /organism="Rattus sp."  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10118"  
 /tissue\_type="liver"  
 /note="788 bp upstream of AvaIII site."

CDS

43..426  
 /codon\_start=1  
 /product="insulin-like growth factor I precursor"  
 /protein\_id="BAA00604.1"  
 /db\_xref="GI:220781"  
 /translation="MSSSHLFYALCLLTFTSSATAGPETLCGAELVDALQFVCGPRG  
 FYFNKPTGYGSSIRRAPQTGIVDECCFRSCDLRRLEMYCAPLKPTKSARSIRAQRHTD  
 MPKTQKEVHLKNTSRGSAGNKTYRM"

mat\_peptide

109..318  
 /product="insulin-like growth factor mature peptide"

variation

606  
 /note="g in [2]; deletion in [1]"

variation

625

```

variation      /note="c in [2]; a in [1]"
               645
polyA_signal   /note="a in [2]; deletion in [1]"
               710..715
polyA_site     731
               /note="polyadenylation site (putative)"
polyA_site     1023
               /note="polyadenylation site (putative)"
BASE COUNT     301 a      256 c      225 g      270 t
ORIGIN

```

```

Query Match          74.7%;  Score 402.6;  DB 10;  Length 1052;
Best Local Similarity 88.5%;  Pred. No. 2.7e-114;
Matches 470;  Conservative 0;  Mismatches 9;  Indels 52;  Gaps 1;

```

```

Qy      1 GGACCAGAGACCCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTCGTGTGTGGACCA 60
      |||
Db      109 GGACCAGAGACCCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTCGTGTGTGGACCA 168

Qy      61 AGGGGCTTTTACTTCAACAAGCCACAGTCTATGGCTCCAGCATTCCGAGGGGCACCACAG 120
      |||
Db      169 AGGGGCTTTTACTTCAACAAGCCACAGGCTATGGCTCCAGCATTCCGAGGGGCACCACAG 228

Qy      121 ACGGGCATTGTGGATGAGTGTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 180
      |||
Db      229 ACGGGCATTGTGGATGAGTGTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 288

Qy      181 TGTGTCCGCTGCAAGCCTACAAAGTCAGCTCGTTCCATCCGGGCCCAGCGCCCACTGAC 240
      |||
Db      289 TGTGCTCCGCTGAAGCCTACAAAGTCAGCTCGTTCCATCCGGGCCCAGCGCCCACTGAC 348

Qy      241 ATGCCCAAGACTCAGAAGTCCCAGCCCCCTATCGACACACAAGAAAAGGAAGCTGCAAAGG 300
      |||
Db      349 ATGCCCAAGACTCAG----- 363

Qy      301 AGAAGGAAAGGAAGTACACTTGAAGAACACAAGTAGAGGAAGTGCAGGAAACAAGACCTA 360
      |||
Db      364 -----AAGGAAGTACACTTGAAGAACACAAGTAGAGGAAGTGCAGGAAACAAGACCTA 416

Qy      361 CAGAATGTAGGAGGAGCCTCCCGAGGAACAGAAAATGCCACGTACCGCAAGATCCTTTG 420
      |||
Db      417 CAGAATGTAGGAGGAGCCTCCCGAGGAACAGAAAATGCCACGTACCGCAAGATCCTTTG 476

Qy      421 CTGCTTGAGCAACCTGCAAAACATCGGAACACCTGCCAAATATCAATAATGAGTTCAATA 480
      |||
Db      477 CTGCTTGAGCAACCTGCAAAACATCGGAACACCTGCCAAATATCAATAATGAGTTCAATA 536

Qy      481 TCATTTTCAGAGATGGGCATTTCCCTCAATGAAATACACAAGTAAACATTCC 531
      |||
Db      537 CCATTTTCAGAGATGGGCATTTCCCTCAATGAAATACACAAGTAAACATTCC 587

```

```

RESULT 12
RNMGJIL
ID  RNMGJIL      standard; RNA; ROD; 513 BP.
XX
AC  M17714;

```

XX  
 SV M17714.1  
 XX  
 DT 01-OCT-1996 (Rel. 49, Created)  
 DT 04-MAR-2000 (Rel. 63, Last updated, Version 2)  
 XX  
 DE Rat insulin-like growth factor-I mRNA, 3' end.  
 XX  
 KW insulin-like growth factor.  
 XX  
 OS Rattus norvegicus (Norway rat)  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;  
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 XX  
 RN [1]  
 RP 1-513  
 RX MEDLINE; 87246437.  
 RX PUBMED; 3595538.  
 RA Murphy L.J., Bell G.I., Duckworth M.L., Friesen H.G.;  
 RT "Identification, characterization, and regulation of a rat complementary  
 RT deoxyribonucleic acid which encodes insulin-like growth factor-I";  
 RL Endocrinology 121(2):684-691(1987).  
 XX  
 DR GOA; P08025.  
 DR SWISS-PROT; P08025; IGFA\_RAT.  
 XX  
 FH Key Location/Qualifiers  
 FH  
 FT source 1. .513  
 FT /db\_xref="taxon:10116"  
 FT /organism="Rattus norvegicus"  
 FT mRNA <1. .513  
 FT /note="ILGF-I mRNA"  
 FT sig\_peptide <1. .9  
 FT /note="insulin-like growth factor-I signal peptide"  
 FT CDS <1. .327  
 FT /codon\_start=1  
 FT /db\_xref="GOA:P08025"  
 FT /db\_xref="SWISS-PROT:P08025"  
 FT /note="insulin-like growth factor-I precursor"  
 FT /protein\_id="AAA41227.1"  
 FT /translation="ATAGPETLCGAELVDALQFVCGPRGFYFNKPTGYGSSIRRAPQTG  
 FT IVDECCFRSCDLRRLEMYCAPLKPTKSARSIRAQRHTDMPKTQKEVHLKNTSRGSAGNK  
 FT TYRM"  
 FT mat\_peptide 10. .324  
 FT /note="insulin-like growth factor-I"  
 XX  
 SQ Sequence 513 BP; 143 A; 128 C; 135 G; 107 T; 0 other;

Query Match 74.4%; Score 401; DB 26; Length 513;  
 Best Local Similarity 88.3%; Pred. No. 7.5e-114;  
 Matches 469; Conservative 0; Mismatches 10; Indels 52; Gaps 1;

Qy 1 GGACCAGAGACCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTCGTGTGTGGACCA 60  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 10 GGACCAGAGACCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTCGTGTGTGGACCA 69

Qy 61 AGGGGCTTTTACTTCAACAAGCCACAGTCTATGGCTCCAGCATTTCGGAGGGCACCACAG 120  
 |||  
 Db 70 AGGGGCTTTTACTTCAACAAGCCACAGGCTATGGCTCCAGCATTTCGGAGGGCACCACAG 129

Qy 121 ACGGGCATTGTGGATGAGTGTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 180  
 |||  
 Db 130 ACGGGCATTGTGGATGAGTGTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 189

Qy 181 TGTGTCCGCTGCAAGCCTACAAAGTCAGCTCGTTCCATCCGGGCCCAGCGCCACACTGAC 240  
 |||  
 Db 190 TGTGCTCCGCTGAAGCCTACAAAGTCAGCTCGTTCCATCCGGGCCCAGCGCCACACTGAC 249

Qy 241 ATGCCCAAGACTCAGAAGTCCCAGCCCCTATCGACACACAAGAAAAGGAAGCTGCAAAGG 300  
 |||  
 Db 250 ATGCCCAAGACTCAG----- 264

Qy 301 AGAAGGAAAGGAAGTACACTTGAAGAACACAAGTAGAGGAAGTGCAGGAAACAAGACCTA 360  
 |||  
 Db 265 -----AAGGAAGTACACTTGAAGAACACAAGTAGAGGAAGTGCAGGAAACAAGACCTA 317

Qy 361 CAGAATGTAGGAGGAGCCTCCCGAGGAACAGAAAATGCCACGTCAACGCAAGATCCTTTG 420  
 |||  
 Db 318 CAGAATGTAGGAGGAGCCTCCCGAGGAACAGAAAATGCCACGTCAACGCAAGATCCTTTG 377

Qy 421 CTGCTTGAGCAACCTGCAAAACATCGGAACACCTGCCAAATATCAATAATGAGTTCAATA 480  
 |||  
 Db 378 CTGCTTGAGGAACCTGCAAAACATCGGAACACCTGCCAAATATCAATAATGAGTTCAATA 437

Qy 481 TCATTTTCAGAGATGGGCATTTCCCTCAATGAAATACACAAGTAAACATTCC 531  
 |||  
 Db 438 CCATTTTCAGAGATGGGCATTTCCCTCAATGAAATACACAAGTAAACATTCC 488

# RESULT 13

## RNIGFIR

LOCUS RNIGFIR 521 bp mRNA linear ROD 31-MAR-2003

DEFINITION Rattus norvegicus mRNA for insulin-like growth factor1 (IGF-1).

ACCESSION X06043 M17714

VERSION X06043.1 GI:56434

KEYWORDS IGF-1; insulin-like growth factor I.

SOURCE Rattus norvegicus (Norway rat)

ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

REFERENCE 1 (bases 1 to 521)

AUTHORS Murphy,L.J., Bell,G.I., Duckworth,M.L. and Friesen,H.G.

TITLE Identification, characterization, and regulation of a rat complementary deoxyribonucleic acid which encodes insulin-like growth factor-I

JOURNAL Endocrinology 121 (2), 684-691 (1987)

MEDLINE 87246437

PUBMED 3595538

COMMENT On Apr 1, 2003 this sequence version replaced gi:204324.

Data kindly reviewed (11-FEB-88) by Murphy L. J.

FEATURES Location/Qualifiers

source 1. 521

```

/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
CDS      <1..327
         /note="(AA -3 to 105)"
         /codon_start=1
         /product="IGF-IA precursor"
         /protein_id="CAA29436.1"
         /db_xref="GI:56435"
         /db_xref="GOA:P08025"
         /db_xref="SWISS-PROT:P08025"
         /translation="ATAGPETLCGAE LVDALQFVCGPRGFYFNKPTGYGSSIRRAPQT
GIVDECCFRSCDLRRLEMYCAPLKPTKSARSIRAQRHTDMPKTQKEVHLKNTSRGSAG
NKTYRM"
sig_peptide <1..9
         /note="insulin-like growth factor-I signal peptide"
mat_peptide 10..324
         /product="insulin-like growth factor-I"
misc_feature 10..96
         /note="B domain (AA 1-29)"
misc_feature 97..132
         /note="C domain (AA 30-41)"
misc_feature 133..195
         /note="A domain (AA 42-62)"
misc_feature 196..219
         /note="D domain (AA 63-70)"
BASE COUNT      151 a      128 c      135 g      107 t
ORIGIN

```

```

Query Match          74.4%;  Score 401;  DB 10;  Length 521;
Best Local Similarity 88.3%;  Pred. No. 7.5e-114;
Matches 469;  Conservative 0;  Mismatches 10;  Indels 52;  Gaps 1;

```

```

Qy      1 GGACCAGAGACCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTCGTGTGTGGACCA 60
      |||
Db      10 GGACCAGAGACCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTCGTGTGTGGACCA 69
      |||
Qy      61 AGGGGCTTTTACTTCAACAAGCCCACAGTCTATGGCTCCAGCATTTCGGAGGGGCACCAACAG 120
      |||
Db      70 AGGGGCTTTTACTTCAACAAGCCCACAGGCTATGGCTCCAGCATTTCGGAGGGGCACCAACAG 129
      |||
Qy      121 ACGGGCATTGTGGATGAGTGTGTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 180
      |||
Db      130 ACGGGCATTGTGGATGAGTGTGTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 189
      |||
Qy      181 TGTGTCCGCTGCAAGCCTACAAAGTCAGCTCGTTCCATCCGGGGCCAGCGCCCACTGAC 240
      |||
Db      190 TGTGCTCCGCTGAAGCCTACAAAGTCAGCTCGTTCCATCCGGGGCCAGCGCCCACTGAC 249
      |||
Qy      241 ATGCCCAAGACTCAGAAGTCCCAGCCCCATATCGACACACAAGAAAAGGAAGCTGCAAAGG 300
      |||
Db      250 ATGCCCAAGACTCAG----- 264
      |||
Qy      301 AGAAGGAAAGGAAGTACACTTGAAGAACACAAGTAGAGGAAGTGCAGGAAACAAGACCTA 360
      |||
Db      265 -----AAGGAAGTACACTTGAAGAACACAAGTAGAGGAAGTGCAGGAAACAAGACCTA 317

```

Qy 361 CAGAATGTAGGAGGAGCCTCCCGAGGAACAGAAAATGCCACGTCACCGCAAGATCCTTTG 420  
 |||  
 Db 318 CAGAATGTAGGAGGAGCCTCCCGAGGAACAGAAAATGCCACGTCACCGCAAGATCCTTTG 377  
 Qy 421 CTGCTTGAGCAACCTGCAAAACATCGGAACACCTGCCAAATATCAATAATGAGTTCAATA 480  
 |||  
 Db 378 CTGCTTGAGGAACCTGCAAAACATCGGAACACCTGCCAAATATCAATAATGAGTTCAATA 437  
 Qy 481 TCATTTTCAGAGATGGGCATTTCCCTCAATGAAATACACAAGTAAACATTCC 531  
 |||  
 Db 438 CCATTTTCAGAGATGGGCATTTCCCTCAATGAAATACACAAGTAAACATTCC 488

# RESULT 14

## RATIGFIAA

LOCUS RATIGFIAA 826 bp mRNA linear ROD 27-APR-1993

DEFINITION Rat insulin-like growth factor I (IGF-I) mRNA, complete cds.

ACCESSION M17335

VERSION M17335.1 GI:204751

KEYWORDS insulin-like growth factor I.

SOURCE Rattus norvegicus (Norway rat)

ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.

REFERENCE 1 (bases 1 to 826)

AUTHORS Casella,S.J., Smith,E.P., van Wyk,J.J., Joseph,D.R., Hynes,M.A.,  
 Hoyt,E.C. and Lund,P.K.

TITLE Isolation of rat testis cDNAs encoding an insulin-like growth  
 factor I precursor

JOURNAL DNA 6 (4), 325-330 (1987)

MEDLINE 88003970

PUBMED 3652906

COMMENT Original source text: Rat (Sprague Dawley) testis, cDNA to mRNA,  
 clone P2.

FEATURES Location/Qualifiers

source 1..826  
 /organism="Rattus norvegicus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10116"

CDS 259..642  
 /note="insulin like growth factor I"  
 /codon\_start=1  
 /protein\_id="AAA41386.1"  
 /db\_xref="GI:204752"  
 /translation="MSSSHLFYLLCLLTFTSSATAGPETLCAELVDALQFVCGPRG  
 FYFNKPTGYGSSIRRAPQTGIVDECCFRSCDLRRLEMYCAPLKPTKSARSIRAQRHTD  
 MPKTQKEVHLKNTSRGSAGNKTYRM"

BASE COUNT 223 a 214 c 181 g 208 t

ORIGIN

Query Match 71.8%; Score 387; DB 10; Length 826;

Best Local Similarity 88.0%; Pred. No. 1.9e-109;

Matches 455; Conservative 0; Mismatches 10; Indels 52; Gaps 1;

Qy 1 GGACCAGAGACCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTCGTGTGTGGACCA 60  
 |||



```

Db      325 GGACCAGAGACCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAATTCGTGTGTGGACCA 384
Qy      61 AGGGGCTTTTACTTCAACAAGCCACAGTCTATGGCTCCAGCATTTCGGAGGGCACCACAG 120
      |||||
Db      385 AGGGGCTTTTACTTCAACAAGCCACAGGCTATGGCTCCAGCATTTCGGAGGGCACCACAG 444
Qy      121 ACGGGCATTGTGGATGAGTGTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 180
      |||||
Db      445 ACGGGCATTGTGGATGAGTGTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 504
Qy      181 TGTGTCCGCTGCAAGCCTACAAAGTCAGCTCGTTCCATCCGGGCCCAGCGCCACACTGAC 240
      |||||
Db      505 TGTGCTCCGCTGAAGCCTACAAAGTCAGCTCGTTCCATCCGGGCCCAGCGCCACACTGAC 564
Qy      241 ATGCCCAAGACTCAGAAGTCCCAGCCCCTATCGACACACAAGAAAAGGAAGCTGCAAAGG 300
      |||||
Db      565 ATGCCCAAGACTCAG----- 579
Qy      301 AGAAGGAAAGGAAGTACACTTGAAGAACACAAGTAGAGGAAGTGCAGGAAACAAGACCTA 360
      |||||
Db      580 -----AAGGAAGTACACTTGAAGAACACAAGTAGAGGAAGTGCAGGAAACAAGACCTA 632
Qy      361 CAGAATGTAGGAGGAGCCTCCCGAGGAACAGAAAATGCCACGTACCCGCAAGATCCTTTG 420
      |||||
Db      633 CAGAATGTAGGAGGAGCCTCCCGAGGAACAGAAAATGCCACGTACCCGCAAGATCCTTTG 692
Qy      421 CTGCTTGAGCAACCTGCAAAACATCGGAACACCTGCCAAATATCAATAATGAGTTCAATA 480
      |||||
Db      693 CTGCTTGAGCAACCTGCAAAACATCGGAACACCTGCCAAATATCAATAATGAGTTCAATA 752
Qy      481 TCATTTTCAGAGATGGGCATTTCCCTCAATGAAATACA 517
      |||||
Db      753 CCATTTTCAGAGATGGGCATTTCCCTCAATGAAATACA 789

```

# RESULT 15

## MMIGFIAR

LOCUS MMIGFIAR 696 bp mRNA linear ROD 21-MAR-1995

DEFINITION Mouse mRNA for preproinsulin-like growth factor IA.

ACCESSION X04480

VERSION X04480.1 GI:51801

KEYWORDS growth factor; insulin-like growth factor IA; preproinsulin-like growth factor IA; signal peptide.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 696)

AUTHORS Bell,G.I., Stempien,M.M., Fong,N.M. and Rall,L.B.

TITLE Sequences of liver cDNAs encoding two different mouse insulin-like growth factor I precursors

JOURNAL Nucleic Acids Res. 14 (20), 7873-7882 (1986)

MEDLINE 87040760

PUBMED 3774549

COMMENT The B-domain of IGF-I comprises residues 1-29 (position 219-305), the C-domain residues 30-41 (position 306-341), the A-domain residues 42-62 (position 342-404) and the D-domain residues 63-70

```

                (position 405-428).
FEATURES             Location/Qualifiers
     source            1..696
                        /organism="Mus musculus"
                        /mol_type="mRNA"
                        /strain="Balb/c"
                        /db_xref="taxon:10090"
                        /clone="migfl-2"
                        /tissue_type="liver"
     misc_feature       123..125
                        /note="pot. translation start"
     CDS                153..536
                        /note="put. preproIGF-IA (AA -22 to 105)"
                        /codon_start=1
                        /protein_id="CAA28168.1"
                        /db_xref="GI:51802"
                        /db_xref="MGI:96432"
                        /db_xref="SWISS-PROT:P05017"
                        /translation="MSSSHLFYLA LCLLTFTSSTTAGPETLCGAELVDALQFVCGPRG
FYFNKPTGYGSSIRRAPQTGIVDECCFRSCDLRRLEMYCAPLKPTKAARSIRAQRHTD
MPKTQKEVHLKNTSRGSAGNKTYRM"
     sig_peptide        153..218
                        /note="signal peptide (AA -22 to -1)"
     mat_peptide        219..428
                        /product="mature IGF-IA (AA 1-70)"
     misc_feature       429..533
                        /note="COOH-terminal peptide (E domain) (AA 71 to 105)"
     polyA_site         696
                        /note="polyA site"
BASE COUNT           201 a    200 c    148 g    147 t
ORIGIN

```

```

Query Match          67.1%;  Score 361.6;  DB 10;  Length 696;
Best Local Similarity 83.8%;  Pred. No. 1.6e-101;
Matches 444;  Conservative 0;  Mismatches 34;  Indels 52;  Gaps 1;

```

```

Qy      1 GGACCAGAGACCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTCGTGTGTGGACCA 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      219 GGACCAGAGACCCTTTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGACCG 278

Qy      61 AGGGGCTTTTACTTCAACAAGCCACAGTCTATGGCTCCAGCATTCCGAGGGGCACCACAG 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      279 AGGGGCTTTTACTTCAACAAGCCACAGGCTATGGCTCCAGCATTCCGAGGGGCACCTCAG 338

Qy      121 ACGGGCATTGTGGATGAGTGTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 180
        || ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      339 ACAGGCATTGTGGATGAGTGTGCTTCCGGAGCTGTGATCTGAGGAGACTGGAGATGTAC 398

Qy      181 TGTGTCCGCTGCAAGCCTACAAAGTCAGCTCGTTCCATCCGGGCCCAGCGCCACACTGAC 240
        |||| || |||||||||||| |||| || || ||||| ||||||||||||||||
Db      399 TGTGCCCCACTGAAGCCTACAAAAGCAGCCGCTCTATCCGTGCCCAGCGCCACACTGAC 458

Qy      241 ATGCCCAAGACTCAGAAGTCCCAGCCCCTATCGACACACAAGAAAAGGAAGCTGCAAAGG 300
        ||||||||||||||||
Db      459 ATGCCCAAGACTCAG----- 473

Qy      301 AGAAGGAAAGGAAGTACACTTGAAGAACACAAGTAGAGGAAGTGCAGGAAACAAGACCTA 360

```

```

          |||||
Db      474 -----AAGGAAGTACATTTGAAGAACACAAGTAGAGGAAGTGCAGGAAACAAGACCTA 526
          |||||
Qy      361 CAGAATGTAGGAGGAGCCTCCCGAGGAACAGAAAATGCCACGTCACCGCAAGATCCTTTG 420
          |||||
Db      527 CAGAATGTAGGAGGAGCCTCCCGAGGAGCAGAAAATGCCACATCACCGCAGGATCCTTTG 586
          |||||
Qy      421 CTGCTTGAGCAACCTGCAAAACATCGGAACACCTGCCAAATATCAATAATGAGTTCAATA 480
          |||||
Db      587 CTGCTTGAGCAACCTGCAAAACATCGAAACACCTACCAAATAACAATAATAAGTCCAATA 646
          |||||
Qy      481 TCATTTTCAGAGATGGGCATTTCCCTCAATGAAATACACAAGTAAACATTTC 530
          |||||
Db      647 ACATTACAAAGATGGGCATTTCCCCCAATGAAATATACAAGTAAACATTTC 696

```

Search completed: December 13, 2003, 09:27:34  
Job time : 2410.26 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 13, 2003, 02:35:18 ; Search time 216.419 Seconds  
(without alignments)  
6723.048 Million cell updates/sec

Title: US-09-852-261-3  
Perfect score: 539  
Sequence: 1 ggaccagagaccctttgcgg.....agtaaacattcccgaattc 539

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_19Jun03:\*

- 1: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:\*
- 2: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:\*
- 3: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:\*
- 4: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:\*
- 5: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:\*
- 6: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:\*
- 7: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:\*
- 8: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:\*
- 9: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:\*
- 10: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:\*
- 11: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:\*
- 12: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:\*
- 13: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:\*
- 14: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:\*
- 15: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:\*
- 16: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:\*
- 17: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:\*
- 18: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:\*
- 19: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:\*
- 20: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:\*
- 21: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:\*
- 22: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:\*
- 23: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:\*
- 24: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:\*
- 25: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	539	100.0	539	22	AAD06399	Rat IGF-I isoform
2	539	100.0	539	24	AAS16878	Rat mechano-growth
3	453.8	84.2	651	25	ABV76185	Mouse insulin-like
4	425	78.8	487	22	AAD06404	Rat liver-type IGF
5	425	78.8	487	24	AAS16883	Rat insulin-like g
6	402.6	74.7	1052	20	AAX27498	Rat liver form of
7	356.8	66.2	523	22	AAD06400	Rabbit IGF-I isofo
8	356.8	66.2	523	24	AAS16879	Rabbit mechano-gro
9	356.8	66.2	553	18	AAT84893	Rabbit insulin lik
10	325.2	60.3	517	22	AAD06398	Human IGF-I isofo
11	325.2	60.3	517	24	AAS16877	Human mechano-grow
12	299.6	55.6	671	24	ABT09479	Phase-1 Rat CT gen
13	274.2	50.9	818	8	AAN70436	Sequence encoding
14	274.2	50.9	7260	24	ABT11091	Human breast cance
15	274.2	50.9	7260	24	ABK84583	Human cDNA differe
16	274.2	50.9	7260	24	ABN97244	Gene #3742 used to
17	274.2	50.9	7260	24	ABK64812	Human benign prost
18	274.2	50.9	7260	24	ABK35504	Human endometrial
19	274.2	50.9	7260	24	ABK35561	Gene IGF1 differen
20	272.6	50.6	777	18	AAT84894	Human insulin like
21	271	50.3	622	7	AAN60490	Human prepro-somat
22	262	48.6	471	22	AAD06405	Rabbit liver-type
23	262	48.6	471	24	AAS16884	Rabbit insulin-lik
24	231	42.9	1136	8	AAN70435	Sequence encoding
25	230.2	42.7	978	14	AAQ47804	Sequence encoding
26	224	41.6	612	22	AAS14695	Human cDNA encodin
27	224	41.6	612	25	ABZ83309	Toxicologically re
28	222	41.2	3599	19	AAV50428	Plasmid pIG0552 lo
29	222	41.2	3599	19	AAV40796	Actual sequence of
30	222	41.2	3600	19	AAV50427	Plasmid pIG0552 up
31	222	41.2	3600	19	AAV40795	Expected sequence
32	222	41.2	5707	20	AAX88055	Plasmid pIG0335 DN
33	222	41.2	6345	20	AAX88054	Plasmid pIG0100A D
34	213.8	39.7	286	25	ABV76186	Mouse insulin-like
35	206.8	38.4	317	24	AAS16882	Human insulin-like
36	206.8	38.4	318	22	AAD06403	Human liver-type I
37	206.8	38.4	462	19	AAV50426	Human IGF-1 encodi
38	206.8	38.4	462	19	AAV40794	Human IGF-I coding
39	206.8	38.4	462	24	ABZ35734	Human IGF1 polynuc
40	206.8	38.4	462	24	ABX09977	Human IGF1 DNA fra
41	206.8	38.4	462	24	ABV78158	Human IGF1 DNA SEQ
42	206.8	38.4	462	24	ABL91699	Human polynucleoti
43	168.4	31.2	210	24	AAD45568	Human insulin-like
44	168.4	31.2	210	24	AAD44955	Human insulin grow
45	168.4	31.2	210	24	ABA03146	Native mature IGF-

# ALIGNMENTS

RESULT 1

AAD06399

ID AAD06399 standard; cDNA; 539 BP.

XX

AC AAD06399;

XX

DT 10-AUG-2001 (first entry)

XX

DE Rat IGF-I isoform mechano-growth factor (MGF) cDNA.

XX

KW Rat; IGF-I isoform; Insulin-like Growth Factor-I; MGF;

KW mechano-growth factor; neurological disorder; neurodegenerative disorder;

KW amyotrophic lateral sclerosis; spinal muscular atrophy; muscular atrophy;

KW poliomyelitis; post-polio syndrome; toxin; motoneurone disorder;

KW nerve damage; autosomal muscular dystrophy; diabetic neuropathy;

KW sex-linked muscular dystrophy; peripheral neuropathy;

KW Alzheimer's disease; Parkinson's disease; ss.

XX

OS Rattus sp.

XX

FH Key Location/Qualifiers

FT CDS 1..336

FT /\*tag= a

FT /product= "Mechano-growth factor (MGF)"

FT /note= "This region comprises exons 3-6. The CDS does not include start codon"

FT /partial

XX

PN WO200136483-A1.

XX

PD 25-MAY-2001.

XX

PF 15-NOV-2000; 2000WO-GB04354.

XX

PR 15-NOV-1999; 99GB-0026968.

XX

PA (UNLO ) UNIV COLLEGE LONDON.

XX

PI Goldspink G, Johnson I;

XX

DR WPI; 2001-355620/37.

DR P-PSDB; AAE02448.

XX

PT Use of mechano-growth factor, an isoform of Insulin-like Growth

PT Factor-I, capable of reducing motoneurone loss, in the manufacture of a

PT medicament for the treatment of neurological disorder -

XX

PS Claim 4; Page 51-52; 66pp; English.

XX

CC The present invention relates to use of mechano-growth factor (MGF),

CC an Insulin-like Growth Factor-I (IGF-I) isoform in the manufacture of a

CC medicament for the treatment of neurological disorder. The MGF is capable

CC of reducing motoneurone loss by 20% or greater in response to nerve

CC avulsion, and effects motoneurone rescue, preferably adult motoneurone

CC rescue. The MGF polynucleotide and polypeptide are useful in the

CC manufacture of a medicament for the treatment of a neurological disorder,

CC including a disorder of motoneurons and/or neurodegenerative disorder,  
 CC e.g., amyotrophic lateral sclerosis, spinal muscular atrophy, progressive  
 CC spinal muscular atrophy, infantile or juvenile muscular atrophy,  
 CC poliomyelitis or post-polio syndrome, a disorder caused by exposure to a  
 CC toxin, motoneuron trauma, a motoneuron lesion or nerve damage, an  
 CC injury that affects motoneurons, motoneuron loss associated with aging,  
 CC autosomal or sex-linked muscular dystrophy, diabetic neuropathy,  
 CC peripheral neuropathies, Alzheimer's disease and Parkinson's disease.  
 CC The present sequence is rat IGF-I isoform MGF cDNA. MGF is a muscle  
 CC isoform having extracellular (Ec) domain, hence also referred as  
 CC IGF-I-Ec. The MGF protein comprises amino acid sequences encoded by  
 CC nucleic acid sequence of IGF-I exons 4, 5 and 6 in the reading frame  
 CC of MGF.

XX

SQ Sequence 539 BP; 161 A; 136 C; 139 G; 103 T; 0 other;

Query Match 100.0%; Score 539; DB 22; Length 539;  
 Best Local Similarity 100.0%; Pred. No. 1.9e-142;  
 Matches 539; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	GGACCAGAGACCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTCGTGTGTGGACCA	60
Db	1	GGACCAGAGACCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTCGTGTGTGGACCA	60
Qy	61	AGGGGCTTTTACTTCAACAAGCCACAGTCTATGGCTCCAGCATTCGGAGGGCACCACAG	120
Db	61	AGGGGCTTTTACTTCAACAAGCCACAGTCTATGGCTCCAGCATTCGGAGGGCACCACAG	120
Qy	121	ACGGGCATTGTGGATGAGTGTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC	180
Db	121	ACGGGCATTGTGGATGAGTGTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC	180
Qy	181	TGTGTCCGCTGCAAGCCTACAAAGTCAGCTCGTTCCATCCGGGCCAGCGCCACACTGAC	240
Db	181	TGTGTCCGCTGCAAGCCTACAAAGTCAGCTCGTTCCATCCGGGCCAGCGCCACACTGAC	240
Qy	241	ATGCCCAAGACTCAGAAGTCCCAGCCCCCTATCGACACACAAGAAAAGGAAGCTGCAAAGG	300
Db	241	ATGCCCAAGACTCAGAAGTCCCAGCCCCCTATCGACACACAAGAAAAGGAAGCTGCAAAGG	300
Qy	301	AGAAGGAAAGGAAGTACACTTGAAGAACACAAGTAGAGGAAGTCAGGAAACAAGACCTA	360
Db	301	AGAAGGAAAGGAAGTACACTTGAAGAACACAAGTAGAGGAAGTCAGGAAACAAGACCTA	360
Qy	361	CAGAATGTAGGAGGAGCCTCCCGAGGAACAGAAAATGCCACGTACCCGCAAGATCCTTTG	420
Db	361	CAGAATGTAGGAGGAGCCTCCCGAGGAACAGAAAATGCCACGTACCCGCAAGATCCTTTG	420
Qy	421	CTGCTTGAGCAACCTGCAAAACATCGGAACACCTGCCAAATATCAATAATGAGTTCAATA	480
Db	421	CTGCTTGAGCAACCTGCAAAACATCGGAACACCTGCCAAATATCAATAATGAGTTCAATA	480
Qy	481	TCATTTTCAGAGATGGGCATTTCCCTCAATGAAATACACAAGTAAACATTCCCCGGAATTC	539
Db	481	TCATTTTCAGAGATGGGCATTTCCCTCAATGAAATACACAAGTAAACATTCCCCGGAATTC	539

RESULT 2

AAS16878

ID AAS16878 standard; cDNA; 539 BP.

XX

AC AAS16878;

XX

DT 25-FEB-2002 (first entry)

XX

DE Rat mechano-growth factor (MGF) cDNA.

XX

KW Rat; mechano-growth factor; insulin-like growth factor I; IGF-I; MGF;  
 KW neuroprotective; nerve damage; peripheral nervous system; nerve severing;  
 KW muscle; neurological disorder; motoneuron loss; motorneuron disorder; ss;  
 KW nerve avulsion.

XX

OS Rattus sp.

XX

FH Key Location/Qualifiers

FT CDS 1..336

FT /\*tag= a

FT /product= "Rat MGF"

FT /partial

FT /note= "No start codon"

FT exon 1..75

FT /\*tag= b

FT /number= exon 3

FT exon 76..258

FT /\*tag= c

FT /number= exon 4

FT exon 259..309

FT /\*tag= d

FT /number= exon 5

FT exon 310..333

FT /\*tag= e

FT /number= exon 6

XX

PN WO200185781-A2.

XX

PD 15-NOV-2001.

XX

PF 10-MAY-2001; 2001WO-GB02054.

XX

PR 10-MAY-2000; 2000GB-0011278.

XX

PA (UNLO ) UNIV COLLEGE LONDON.

PA (EGRI-) EAST GRINSTEAD MEDICAL RES TRUST.

XX

PI Goldspink G, Terenghi G;

XX

DR WPI; 2002-055585/07.

DR P-PSDB; AAU10560.

XX

PT Use of insulin-like growth factor I (IGF-I) isoform known as  
 PT mechano-growth factor which is encoded by IGF-I exons 4,5,6 and has  
 PT ability to reduce motoneuron loss in response to nerve avulsion, to  
 PT treat nerve damage -

XX



PS Disclosure; Fig 6; 65pp; English.

XX

CC The invention relates to the use of an insulin-like growth factor I  
CC (IGF-I) isoform, known as mechano-growth factor (MGF), in the manufacture  
CC of a medicament for treating nerve damage in the peripheral nervous  
CC system, or for treating nerve damage by localising MGF at the site of  
CC damage. The nerve damage may include severing of a nerve. The treatment  
CC may be combined with another treatment (such as a polypeptide growth  
CC factor other than MGF) that prevents or diminishes degeneration of the  
CC target organ (for example, muscle) which the damaged nerve innervates,  
CC whereby the treatment of the muscle with MGF or a polynucleotide encoding  
CC MGF prevents or diminishes degeneration. The method is useful for  
CC treating neurological disorders, preferably motorneuron disorders. These  
CC methods can reduce motoneuron loss by 20% or greater in response to nerve  
CC avulsion. This sequence represents cDNA encoding the rat MGF.

XX

SQ Sequence 539 BP; 161 A; 136 C; 139 G; 103 T; 0 other;

Query Match 100.0%; Score 539; DB 24; Length 539;  
Best Local Similarity 100.0%; Pred. No. 1.9e-142;  
Matches 539; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 GGACCAGAGACCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTCGTGTGTGGACCA 60
      |||
Db      1 GGACCAGAGACCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTCGTGTGTGGACCA 60

Qy     61 AGGGGCTTTTACTTCAACAAGCCACAGTCTATGGCTCCAGCATTTCGGAGGGCACCACAG 120
      |||
Db     61 AGGGGCTTTTACTTCAACAAGCCACAGTCTATGGCTCCAGCATTTCGGAGGGCACCACAG 120

Qy    121 ACGGGCATTGTGGATGAGTGTGTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 180
      |||
Db    121 ACGGGCATTGTGGATGAGTGTGTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 180

Qy    181 TGTGTCCGCTGCAAGCCTACAAAGTCAGCTCGTTCCATCCGGGCCCAGCGCCCACTGAC 240
      |||
Db    181 TGTGTCCGCTGCAAGCCTACAAAGTCAGCTCGTTCCATCCGGGCCCAGCGCCCACTGAC 240

Qy    241 ATGCCCAAGACTCAGAAGTCCCAGCCCCTATCGACACACAAGAAAAGGAAGCTGCAAAGG 300
      |||
Db    241 ATGCCCAAGACTCAGAAGTCCCAGCCCCTATCGACACACAAGAAAAGGAAGCTGCAAAGG 300

Qy    301 AGAAGGAAAGGAAGTACACTTGAAGAACACAAGTAGAGGAAGTGCAGGAAACAAGACCTA 360
      |||
Db    301 AGAAGGAAAGGAAGTACACTTGAAGAACACAAGTAGAGGAAGTGCAGGAAACAAGACCTA 360

Qy    361 CAGAATGTAGGAGGAGCCTCCCGAGGAACAGAAAATGCCACGTCACCGCAAGATCCTTTG 420
      |||
Db    361 CAGAATGTAGGAGGAGCCTCCCGAGGAACAGAAAATGCCACGTCACCGCAAGATCCTTTG 420

Qy    421 CTGCTTGAGCAACCTGCAAAACATCGGAACACCTGCCAAATATCAATAATGAGTTCAATA 480
      |||
Db    421 CTGCTTGAGCAACCTGCAAAACATCGGAACACCTGCCAAATATCAATAATGAGTTCAATA 480

Qy    481 TCATTTTCAGAGATGGGCATTTCCCTCAATGAAATACACAAGTAAACATTCCCGGAATTC 539
      |||
Db    481 TCATTTTCAGAGATGGGCATTTCCCTCAATGAAATACACAAGTAAACATTCCCGGAATTC 539
```

RESULT 3

ABV76185

ID ABV76185 standard; cDNA; 651 BP.

XX

AC ABV76185;

XX

DT 07-MAR-2003 (first entry)

XX

DE Mouse insulin-like growth factor IB cDNA.

XX

KW Insulin-like growth factor IB; IGF-IB; mouse; mRNA; assay;  
KW nucleic acid detection; gene; ss.

XX

OS Mus musculus.

XX

FH Key Location/Qualifiers

FT CDS 73..474

FT /\*tag= a

FT /product= "IGF-IB"

XX

PN WO200297390-A2.

XX

PD 05-DEC-2002.

XX

PF 31-MAY-2002; 2002WO-SE01056.

XX

PR 01-JUN-2001; 2001SE-0001934.

XX

PA (BIOV-) BIOVITRUM AB.

XX

PI Parrow V, Rosengren L;

XX

DR WPI; 2003-129529/12.

XX

PT Quantitating a target nucleic acid in a sample comprises immobilizing,  
PT on a solid support, a sample comprising a target nucleic acid, and  
PT detecting and quantitating signals generated from the antisense and  
PT sense probes -

XX

PS Example 1; Page 16-17; 18pp; English.

XX

CC The present sequence is that of cDNA encoding murine insulin-like  
CC growth factor 1B (IGF-IB). The cDNA was used in an example of the  
CC method of the invention to generate probes for determination of  
CC IGF-IB RNA. The method comprises a quantitative hybridisation  
CC assay for analysis of mRNA in a target nucleic acid (TNA) sample.  
CC It involves: (i) immobilising the TNA sample on a solid support;  
CC (ii) contacting a labelled antisense probe to a first portion of the  
CC TNA, and a labelled sense probe to a second portion of the TNA;  
CC (iii) detecting and quantitating the signals generated from the  
CC hybridised probes; and (iv) determining the value represented by  
CC the antisense probe signal minus the sense probe signal, the value  
CC being proportional to the amount of mRNA in the TNA sample. In an  
CC example of the method, a cDNA clone containing 60 nucleotides from  
CC exon 2 and 179 nucleotides from exon 3 of the mouse IGF-IB gene was

CC cloned into pGEN-4Z vector. Linearisation of the plasmid with  
 CC EcoRI allowed transcription of a 250-nucleotide antisense probe  
 CC using T7 polymerase. Linearisation with HindIII allowed  
 CC transcription of a sense probe of similar length using SP6  
 CC polymerase (see ABV76186). The probes were purified and used to  
 CC determine IGF-I RNA in mouse hepatocytes and also in rat hepatocytes.

XX

SQ Sequence 651 BP; 193 A; 185 C; 149 G; 124 T; 0 other;

Query Match 84.2%; Score 453.8; DB 25; Length 651;  
 Best Local Similarity 92.8%; Pred. No. 2.5e-118;  
 Matches 476; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

```

Qy      1 GGACCAGAGACCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTCGTGTGTGGACCA 60
      |||
Db      139 GGACCAGAGACCCTTTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGACCG 198

Qy      61 AGGGGCTTTTACTTCAACAAGCCCACAGTCTATGGCTCCAGCATTCCGAGGGGCACACAG 120
      |||
Db      199 AGGGGCTTTTACTTCAACAAGCCCACAGGCTATGGCTCCAGCATTCCGAGGGGCACCTCAG 258

Qy      121 ACGGGCATTGTGGATGAGTGTGTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 180
      ||
Db      259 ACAGGCATTGTGGATGAGTGTGTGCTTCCGGAGCTGTGATCTGAGGAGACTGGAGATGTAC 318

Qy      181 TGTGTCCGCTGCAAGCCTACAAAGTCAGCTCGTTCCATCCGGGGCCAGCGCCCACTGAC 240
      |||
Db      319 TGTGCCCCACTGAAGCCTACAAAGCAGCCCGCTCTATCCGTGCCCAGCGCCCACTGAC 378

Qy      241 ATGCCCAAGACTCAGAAGTCCCAGCCCCATCGACACACAAGAAAAGGAAGCTGCAAAGG 300
      |||
Db      379 ATGCCCAAGACTCAGAAGTCCCCGTCCCTATCGACAAAACAAGAAACGAAGCTGCAAAGG 438

Qy      301 AGAAGGAAAGGAAGTACACTTGAAGAACACAAGTAGAGGAAGTGCAAGAAACAAGACCTA 360
      |||
Db      439 AGAAGGAAAGGAAGTACATTGGAAGAACACAAGTAGAGGAAGTGCAAGAAACAAGACCTA 498

Qy      361 CAGAATGTAGGAGGAGCCTCCCGAGGAACAGAAAATGCCACGTCACCGCAAGATCCTTTG 420
      |||
Db      499 CAGAATGTAGGAGGAGCCTCCCGAGGAGCAGAAAATGCCACATCACCGCAGGATCCTTTG 558

Qy      421 CTGCTTGAGCAACCTGCAAAACATCGGAACACCTGCCAAATATCAATAATGAGTTCAATA 480
      |||
Db      559 CTGCTTGAGCAACCTGCAAAACATCGAAACACCTACCAAATAACAATAATAAGTCCAATA 618

Qy      481 TCATTTTCAGAGATGGGCATTTCCCTCAATGAAA 513
      |||
Db      619 ACATTACAAAGATGGGCATTTCCCCCAATGAAA 651
  
```

# RESULT 4

AAD06404

ID AAD06404 standard; cDNA; 487 BP.

XX

AC AAD06404;

XX

DT 10-AUG-2001 (first entry)

XX  
 DE Rat liver-type IGF-I isoform (L.IGF-I) cDNA.  
 XX  
 KW Rat; IGF-I isoform; Insulin-like Growth Factor-I; MGF;  
 KW mechano-growth factor; neurological disorder; neurodegenerative disorder;  
 KW amyotrophic lateral sclerosis; spinal muscular atrophy; muscular atrophy;  
 KW poliomyelitis; post-polio syndrome; toxin; motoneurone disorder;  
 KW nerve damage; autosomal muscular dystrophy; diabetic neuropathy;  
 KW sex-linked muscular dystrophy; peripheral neuropathy;  
 KW Alzheimer's disease; Parkinson's disease; liver; L.IGF-I; ss.  
 XX  
 OS Rattus sp.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 1..318  
 FT /\*tag= a  
 FT /product= "Liver-type IGF-I isoform (L.IGF-I)"  
 FT /transl\_except= (pos:304..306, aa:Tyr)  
 FT /note= "The translation exception occurs while  
 FT decoding the alternative version of the protein  
 FT (AAE02531). The CDS comprises exons 3, 4 and 6 and does  
 FT not include start codon"  
 FT /partial  
 XX  
 PN WO200136483-A1.  
 XX  
 PD 25-MAY-2001.  
 XX  
 PF 15-NOV-2000; 2000WO-GB04354.  
 XX  
 PR 15-NOV-1999; 99GB-0026968.  
 XX  
 PA (UNLO ) UNIV COLLEGE LONDON.  
 XX  
 PI Goldspink G, Johnson I;  
 XX  
 DR WPI; 2001-355620/37.  
 DR P-PSDB; AAE02451, AAE02531.  
 XX  
 PT Use of mechano-growth factor, an isoform of Insulin-like Growth  
 PT Factor-I, capable of reducing motoneurone loss, in the manufacture of a  
 PT medicament for the treatment of neurological disorder -  
 XX  
 PS Disclosure; Page 57-58; 66pp; English.  
 XX  
 CC The present invention relates to use of mechano-growth factor (MGF),  
 CC an Insulin-like Growth Factor-I (IGF-I) isoform in the manufacture of a  
 CC medicament for the treatment of neurological disorder. The MGF is capable  
 CC of reducing motoneurone loss by 20% or greater in response to nerve  
 CC avulsion, and effects motoneurone rescue, preferably adult motoneurone  
 CC rescue. The MGF polynucleotide and polypeptide are useful in the  
 CC manufacture of a medicament for the treatment of a neurological disorder,  
 CC including a disorder of motoneurons and/or neurodegenerative disorder,  
 CC e.g., amyotrophic lateral sclerosis, spinal muscular atrophy, progressive  
 CC spinal muscular atrophy, infantile or juvenile muscular atrophy,  
 CC poliomyelitis or post-polio syndrome, a disorder caused by exposure to a  
 CC toxin, motoneurone trauma, a motoneurone lesion or nerve damage, an

CC injury that affects motoneurons, motoneurone loss associated with aging,  
 CC autosomal or sex-linked muscular dystrophy, diabetic neuropathy,  
 CC peripheral neuropathies, Alzheimer's disease and Parkinson's disease.  
 CC The present sequence is rat liver-type IGF-I isoform (L.IGF-I) cDNA.  
 CC The L.IGF-I protein comprises amino acid sequences encoded by  
 CC nucleic acid sequence of IGF-I exons 4 and 6.

XX

SQ Sequence 487 BP; 139 A; 123 C; 126 G; 99 T; 0 other;

Query Match 78.8%; Score 425; DB 22; Length 487;  
 Best Local Similarity 90.4%; Pred. No. 3.1e-110;  
 Matches 487; Conservative 0; Mismatches 0; Indels 52; Gaps 1;

```

Qy      1 GGACCAGAGACCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTCGTGTGTGGACCA 60
      |||
Db      1 GGACCAGAGACCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTCGTGTGTGGACCA 60

Qy     61 AGGGGCTTTTACTTCAACAAGCCACAGTCTATGGCTCCAGCATTTCGGAGGGGCACACAG 120
      |||
Db     61 AGGGGCTTTTACTTCAACAAGCCACAGTCTATGGCTCCAGCATTTCGGAGGGGCACACAG 120

Qy    121 ACGGGCATTGTGGATGAGTGTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 180
      |||
Db    121 ACGGGCATTGTGGATGAGTGTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 180

Qy    181 TGTGTCCGCTGCAAGCCTACAAAGTCAGCTCGTTCCATCCGGGCCAGCGCCACACTGAC 240
      |||
Db    181 TGTGTCCGCTGCAAGCCTACAAAGTCAGCTCGTTCCATCCGGGCCAGCGCCACACTGAC 240

Qy    241 ATGCCCAAGACTCAGAAGTCCCAGCCCCTATCGACACACAAGAAAAGGAAGCTGCAAAGG 300
      |||
Db    241 ATGCCCAAGACTCAG----- 255

Qy    301 AGAAGGAAAGGAAGTACACTTGAAGAACACAAGTAGAGGAAGTGCAGGAAACAAGACCTA 360
      |||
Db    256 -----AAGGAAGTACACTTGAAGAACACAAGTAGAGGAAGTGCAGGAAACAAGACCTA 308

Qy    361 CAGAATGTAGGAGGAGCCTCCCGAGGAACAGAAAATGCCACGTCACCGCAAGATCCTTTG 420
      |||
Db    309 CAGAATGTAGGAGGAGCCTCCCGAGGAACAGAAAATGCCACGTCACCGCAAGATCCTTTG 368

Qy    421 CTGCTTGAGCAACCTGCAAAACATCGGAACACCTGCCAAATATCAATAATGAGTTCAATA 480
      |||
Db    369 CTGCTTGAGCAACCTGCAAAACATCGGAACACCTGCCAAATATCAATAATGAGTTCAATA 428

Qy    481 TCATTTTCAGAGATGGGCATTTCCCTCAATGAAATACACAAGTAAACATTCCCGGAATTC 539
      |||
Db    429 TCATTTTCAGAGATGGGCATTTCCCTCAATGAAATACACAAGTAAACATTCCCGGAATTC 487
  
```

RESULT 5

AAS16883

ID AAS16883 standard; cDNA; 487 BP;

XX

AC AAS16883;

XX

DT 25-FEB-2002 (first entry)

XX  
 DE Rat insulin-like growth factor I liver-type isoform (L.IGF-I) cDNA.  
 XX  
 KW Rat; mechano-growth factor; insulin-like growth factor I; IGF-I; MGF;  
 KW neuroprotective; nerve damage; peripheral nervous system; nerve severing;  
 KW muscle; neurological disorder; motoneuron loss; motoneuron disorder; ss;  
 KW nerve avulsion; insulin-like growth factor I liver-type isoform; L.IGF-I;  
 XX  
 OS Rattus sp.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 1..318  
 FT /\*tag= a  
 FT /product= "Rat L.IGF-I"  
 FT /partial  
 FT /note= "No start codon"  
 FT exon 1..75  
 FT /\*tag= b  
 FT /number= exon 3  
 FT exon 76..258  
 FT /\*tag= c  
 FT /number= exon 4  
 FT exon 259..315  
 FT /\*tag= d  
 FT /number= exon 6  
 XX  
 PN WO200185781-A2.  
 XX  
 PD 15-NOV-2001.  
 XX  
 PF 10-MAY-2001; 2001WO-GB02054.  
 XX  
 PR 10-MAY-2000; 2000GB-0011278.  
 XX  
 PA (UNLO ) UNIV COLLEGE LONDON.  
 PA (EGRI-) EAST GRINSTEAD MEDICAL RES TRUST.  
 XX  
 PI Goldspink G, Terenghi G;  
 XX  
 DR WPI; 2002-055585/07.  
 DR P-PSDB; AAU10563.  
 XX  
 PT Use of insulin-like growth factor I (IGF-I) isoform known as  
 PT mechano-growth factor which is encoded by IGF-I exons 4,5,6 and has  
 PT ability to reduce motoneuron loss in response to nerve avulsion, to  
 PT treat nerve damage -  
 XX  
 PS Disclosure; Fig 9; 65pp; English.  
 XX  
 CC The invention relates to the use of an insulin-like growth factor I  
 CC (IGF-I) isoform, known as mechano-growth factor (MGF), in the manufacture  
 CC of a medicament for treating nerve damage in the peripheral nervous  
 CC system, or for treating nerve damage by localising MGF at the site of  
 CC damage. The nerve damage may include severing of a nerve. The treatment  
 CC may be combined with another treatment (such as a polypeptide growth  
 CC factor other than MGF) that prevents or diminishes degeneration of the  
 CC target organ (for example, muscle) which the damaged nerve innervates,

CC whereby the treatment of the muscle with MGF or a polynucleotide encoding  
CC MGF prevents or diminishes degeneration. The method is useful for  
CC treating neurological disorders, preferably motoneuron disorders. These  
CC methods can reduce motoneuron loss by 20% or greater in response to nerve  
CC avulsion. This sequence represents cDNA encoding the rat insulin-like  
CC growth factor I liver-type isoform (L.IGF-I) used in experiments on  
CC motoneuron loss.

XX

SQ Sequence 487 BP; 139 A; 123 C; 126 G; 99 T; 0 other;

Query Match 78.8%; Score 425; DB 24; Length 487;  
Best Local Similarity 90.4%; Pred. No. 3.1e-110;  
Matches 487; Conservative 0; Mismatches 0; Indels 52; Gaps 1;

```
Qy      1 GGACCAGAGACCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTCGTGTGTGGACCA 60
      |||
Db      1 GGACCAGAGACCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTCGTGTGTGGACCA 60

Qy     61 AGGGGCTTTTACTTCAACAAGCCCACAGTCTATGGCTCCAGCATTTCGGAGGGCACCACAG 120
      |||
Db     61 AGGGGCTTTTACTTCAACAAGCCCACAGTCTATGGCTCCAGCATTTCGGAGGGCACCACAG 120

Qy    121 ACGGGCATTGTGGATGAGTGTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 180
      |||
Db    121 ACGGGCATTGTGGATGAGTGTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 180

Qy    181 TGTGTCCGCTGCAAGCCTACAAAGTCAGCTCGTTCCATCCGGGGCCAGCGCCCACTGAC 240
      |||
Db    181 TGTGTCCGCTGCAAGCCTACAAAGTCAGCTCGTTCCATCCGGGGCCAGCGCCCACTGAC 240

Qy    241 ATGCCCAAGACTCAGAAGTCCCAGCCCCCTATCGACACACAAGAAAAGGAAGCTGCAAAGG 300
      |||
Db    241 ATGCCCAAGACTCAG----- 255

Qy    301 AGAAGGAAAGGAAGTACACTTGAAGAACACAAGTAGAGGAAGTGCAGGAAACAAGACCTA 360
      |||
Db    256 -----AAGGAAGTACACTTGAAGAACACAAGTAGAGGAAGTGCAGGAAACAAGACCTA 308

Qy    361 CAGAATGTAGGAGGAGCCTCCCGAGGAACAGAAAATGCCACGTCACCGCAAGATCCTTTG 420
      |||
Db    309 CAGAATGTAGGAGGAGCCTCCCGAGGAACAGAAAATGCCACGTCACCGCAAGATCCTTTG 368

Qy    421 CTGCTTGAGCAACCTGCAAAACATCGGAACACCTGCCAAATATCAATAATGAGTTCAATA 480
      |||
Db    369 CTGCTTGAGCAACCTGCAAAACATCGGAACACCTGCCAAATATCAATAATGAGTTCAATA 428

Qy    481 TCATTTTCAGAGATGGGCATTTCCCTCAATGAAATACACAAGTAAACATTCCCGGAATTC 539
      |||
Db    429 TCATTTTCAGAGATGGGCATTTCCCTCAATGAAATACACAAGTAAACATTCCCGGAATTC 487
```

RESULT 6

AAX27498

ID AAX27498 standard; DNA; 1052 BP.

XX

AC AAX27498;

XX

DT 22-JUN-1999 (first entry)  
 XX  
 DE Rat liver form of IGF-1.  
 XX  
 KW Liver; isoform; rat; insulin-like growth factor; IGF1; vertebrate;  
 KW muscle mass; human; gravity; cosmetic body sculpting; glucose clearance;  
 KW diabetic; ss.  
 XX  
 OS Rattus sp.  
 XX  
 PN WO9910013-A1.  
 XX  
 PD 04-MAR-1999.  
 XX  
 PF 25-AUG-1998; 98WO-US17428.  
 XX  
 PR 25-AUG-1997; 97US-0057201.  
 XX  
 PA (MASS-) MASSACHUSETTS GEN HOSPITAL.  
 PA (UYPE-) UNIV PENNSYLVANIA.  
 XX  
 PI Rosenthal NA, Sweeney HL;  
 XX  
 DR WPI; 1999-190469/16.  
 XX  
 PT Use of modified insulin-like growth factor I - in gene therapy for  
 PT increasing vertebrate muscle strength and mass  
 XX  
 PS Example 1; Fig 4; 46pp; English.  
 XX  
 CC This sequence represents the liver isoform of the rat insulin-like  
 CC growth factor 1 (IGF1) from rats. IGF-1 or a modified or biologically  
 CC active portion can be used for increasing vertebrate muscle mass by  
 CC intramuscular administration, for e.g. (i) preserving or enhancing  
 CC muscle strength in aging humans; (ii) healing injured muscle more  
 CC efficiently/rapidly; (iii) controlling muscle mass during disease and/or  
 CC prolonged stays in reduced gravity; (iv) cosmetic body sculpting; and  
 CC (v) promoting glucose clearance from diabetic muscle tissue.  
 XX  
 SQ Sequence 1052 BP; 301 A; 256 C; 225 G; 270 T; 0 other;

Query Match 74.7%; Score 402.6; DB 20; Length 1052;  
 Best Local Similarity 88.5%; Pred. No. 8.9e-104;  
 Matches 470; Conservative 0; Mismatches 9; Indels 52; Gaps 1;

```

Qy      1 GGACCAGAGACCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTTCGTGTGTGGACCA 60
      |||
Db     109 GGACCAGAGACCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTTCGTGTGTGGACCA 168

Qy     61 AGGGGCTTTTACTTCAACAAGCCCACAGTCTATGGCTCCAGCATTCGGAGGGCACCACAG 120
      |||
Db     169 AGGGGCTTTTACTTCAACAAGCCCACAGGCTATGGCTCCAGCATTCGGAGGGCACCACAG 228

Qy     121 ACGGGCATTGTGGATGAGTGTTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 180
      |||
Db     229 ACGGGCATTGTGGATGAGTGTTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 288
  
```



Qy 181 TGTGTCCGCTGCAAGCCTACAAAGTCAGCTCGTTCCATCCGGGCCAGCGCCACACTGAC 240  
 ||||| | |||||  
 Db 289 TGTGCTCCGCTGAAGCCTACAAAGTCAGCTCGTTCCATCCGGGCCAGCGCCACACTGAC 348  
 Qy 241 ATGCCCCAAGACTCAGAAGTCCCAGCCCCTATCGACACACAAGAAAAGGAAGCTGCAAAGG 300  
 |||||  
 Db 349 ATGCCCCAAGACTCAG----- 363  
 Qy 301 AGAAGGAAAGGAAGTACACTTGAAGAACACAAGTAGAGGAAGTGCAGGAAACAAGACCTA 360  
 |||||  
 Db 364 -----AAGGAAGTACACTTGAAGAACACAAGTAGAGGAAGTGCAGGAAACAAGACCTA 416  
 Qy 361 CAGAATGTAGGAGGAGCCTCCCGAGGAACAGAAAATGCCACGTACCGCAAGATCCTTTG 420  
 |||||  
 Db 417 CAGAATGTAGGAGGAGCCTCCCGAGGAACAGAAAATGCCACGTACCGCAAGATCCTTTG 476  
 Qy 421 CTGCTTGAGCAACCTGCAAAACATCGGAACACCTGCCAAATATCAATAATGAGTTCAATA 480  
 |||||  
 Db 477 CTGCTTGAGCAACCTGCAAAACATCGGAACACCTGCCAAATATCAATAATGAGTTCAATA 536  
 Qy 481 TCATTTTCAGAGATGGGCATTTCCCTCAATGAAATACACAAGTAAACATTCC 531  
 |||||  
 Db 537 CCATTTTCAGAGATGGGCATTTCCCTCAATGAAATACACAAGTAAACATTCC 587

RESULT 7

AAD06400

ID AAD06400 standard; cDNA; 523 BP.

XX

AC AAD06400;

XX

DT 10-AUG-2001 (first entry)

XX

DE Rabbit IGF-I isoform mechano-growth factor (MGF) cDNA.

XX

KW Rabbit; IGF-I isoform; Insulin-like Growth Factor-I; MGF;  
 KW mechano-growth factor; neurological disorder; neurodegenerative disorder;  
 KW amyotrophic lateral sclerosis; spinal muscular atrophy; muscular atrophy;  
 KW poliomyelitis; post-polio syndrome; toxin; motoneurone disorder;  
 KW nerve damage; autosomal muscular dystrophy; diabetic neuropathy;  
 KW sex-linked muscular dystrophy; peripheral neuropathy;  
 KW Alzheimer's disease; Parkinson's disease; ss.

XX

OS Oryctolagus cuniculus.

XX

FH Key Location/Qualifiers

FT CDS 1..336

FT /\*tag= a

FT /product= "Mechano-growth factor (MGF)"

FT /note= "This region comprises exons 3-6. The CDS does  
 FT not include start codon"

FT /partial

XX

PN WO200136483-A1.

XX

PD 25-MAY-2001.

XX

PF 15-NOV-2000; 2000WO-GB04354.

PR 15-NOV-1999; 99GB-0026968.

PA (UNLO ) UNIV COLLEGE LONDON.

PI Goldspink G, Johnson I;

DR WPI; 2001-355620/37.

DR P-PSDB; AAE02449.

PT Use of mechano-growth factor, an isoform of Insulin-like Growth  
PT Factor-I, capable of reducing motoneurone loss, in the manufacture of a  
PT medicament for the treatment of neurological disorder -

XX

PS Claim 4; Page 53-54; 66pp; English.

CC The present invention relates to use of mechano-growth factor (MGF),  
CC an Insulin-like Growth Factor-I (IGF-I) isoform in the manufacture of a  
CC medicament for the treatment of neurological disorder. The MGF is capable  
CC of reducing motoneurone loss by 20% or greater in response to nerve  
CC avulsion, and effects motoneurone rescue, preferably adult motoneurone  
CC rescue. The MGF polynucleotide and polypeptide are useful in the  
CC manufacture of a medicament for the treatment of a neurological disorder,  
CC including a disorder of motoneurons and/or neurodegenerative disorder,  
CC e.g., amyotrophic lateral sclerosis, spinal muscular atrophy, progressive  
CC spinal muscular atrophy, infantile or juvenile muscular atrophy,  
CC poliomyelitis or post-polio syndrome, a disorder caused by exposure to a  
CC toxin, motoneurone trauma, a motoneurone lesion or nerve damage, an  
CC injury that affects motoneurons, motoneurone loss associated with aging,  
CC autosomal or sex-linked muscular dystrophy, diabetic neuropathy,  
CC peripheral neuropathies, Alzheimer's disease and Parkinson's disease.  
CC The present sequence is rabbit IGF-I isoform MGF cDNA. MGF is a muscle  
CC isoform having extracellular (Ec) domain, hence also referred as  
CC IGF-I-Ec. The MGF protein comprises amino acid sequences encoded by  
CC nucleic acid sequence of IGF-I exons 4, 5 and 6 in the reading frame  
CC of MGF.

XX

SQ Sequence 523 BP; 154 A; 129 C; 142 G; 98 T; 0 other;

Query Match 66.2%; Score 356.8; DB 22; Length 523;

Best Local Similarity 82.3%; Pred. No. 6.3e-91;

Matches 436; Conservative 0; Mismatches 87; Indels 7; Gaps 2;

QY            1 GGACCAGAGACCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTCGTGTGTGGACCA 60  
             |||||

Db 1 GGACCGGAGACGCTCTGCGGTGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 60

Qy 61 AGGGGCTTTTACTTCAACAAGCCCAAGTCTATGGCTCCAGCATTCGGAGGGCACACAG 120

Db . 61 AGGGGCTTTTATTTCAACAAGCCACAGGATACGGCTCCAGCAGTCGGAGGGGCACCTCAG 120

QY 121 ACGGGCATTGTGGATGAGTGTTCCTCCGAGACTGTGATCTGAGGAGGCTGGAGATGTAC 180

Db 121 ACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 180

Qy 181 TGTGTCCGCTGCAAGCCTACAAAGTCAGCTCGTTCCATCCGGGCCAGCGCCCACTGAC 240

```

      ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      181 TGTGCACCCCTCAAGCCGGCAAAGGCAGCCCGCTCCGTCCGTGCCAGCGCCACACCGAC 240
Qy      241 ATGCCCAAGACTCAGAAGTCCCAGCCCCCTATCGACACACAAGAAAAGGAAGCTGCAAAGG 300
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      241 ATGCCCAAGACTCAGAAGTATCAGCCTCCATCTACCAACAAGAAAATGAAGTCTCAGAGG 300
Qy      301 AGAAGGAAAGGAAGTACACTTGAAGAACACAAGTAGAGGAAGTGCAGGAAACAAGACCTA 360
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      301 AGAAGGAAAGGAAGTACATTTGAAGAACACAAGTAGAGGGAGTGCAGGAAACAAGAACTA 360
Qy      361 CAGAATGTAGGAGGAGCCTCCCGAGGAACAGAAAATGCCACGTCACCGCAAGATCCTTTG 420
      ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      361 CAGGATGTAGGAAGACCTTCTGAGGAGTGAAGAAGGACAGGCCACCGCAGGACCTTTG 420
Qy      421 CTGCTTGAGCAACCTGCAAAACATCGGAACACCTGCCAAATATCAATAATGAGTTCAATA 480
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      421 CTCTGCACAGTTACCTGTAAACATTGGAATACCGGCCA-----AAAAATAAGTTTGATC 474
Qy      481 TCATTTTCAGAGATGGGCATTTCCCTCAATGAAATACACAAGTAAACATTTC 530
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      475 ACATTTCAAAGAT-GGCATTTCCCCCAATGAAATACACAAGTAAACATTTC 523

```

# RESULT 8

AAS16879

ID AAS16879 standard; cDNA; 523 BP.

XX

AC AAS16879;

XX

DT 25-FEB-2002 (first entry)

XX

DE Rabbit mechano-growth factor (MGF) cDNA.

XX

KW Rabbit; mechano-growth factor; insulin-like growth factor I; IGF-I; MGF;  
 KW neuroprotective; nerve damage; peripheral nervous system; nerve severing;  
 KW muscle; neurological disorder; motoneuron loss; motoneuron disorder; ss;  
 KW nerve avulsion.

XX

OS Oryctolagus cuniculus.

XX

Key	Location/Qualifiers
FT CDS	1..336
FT	/*tag= a
FT	/product= "Rabbit MGF"
FT	/partial
FT	/note= "No start codon"
FT exon	1..76
FT	/*tag= b
FT	/number= 3
FT exon	77..259
FT	/*tag= c
FT	/number= 4
FT exon	260..309
FT	/*tag= d
FT	/number= 5
FT exon	311..333

```

FT          /*tag= e
FT          /number= 6
XX
PN WO200185781-A2.
XX
PD 15-NOV-2001.
XX
PF 10-MAY-2001; 2001WO-GB02054.
XX
PR 10-MAY-2000; 2000GB-0011278.
XX
PA (UNLO ) UNIV COLLEGE LONDON.
PA (EGRI-) EAST GRINSTEAD MEDICAL RES TRUST.
XX
PI Goldspink G, Terenghi G;
XX
DR WPI; 2002-055585/07.
DR P-PSDB; AAU10561.
XX
PT Use of insulin-like growth factor I (IGF-I) isoform known as
PT mechano-growth factor which is encoded by IGF-I exons 4,5,6 and has
PT ability to reduce motoneuron loss in response to nerve avulsion, to
PT treat nerve damage -
XX
PS Disclosure; Fig 7; 65pp; English.
XX
CC The invention relates to the use of an insulin-like growth factor I
CC (IGF-I) isoform, known as mechano-growth factor (MGF), in the manufacture
CC of a medicament for treating nerve damage in the peripheral nervous
CC system, or for treating nerve damage by localising MGF at the site of
CC damage. The nerve damage may include severing of a nerve. The treatment
CC may be combined with another treatment (such as a polypeptide growth
CC factor other than MGF) that prevents or diminishes degeneration of the
CC target organ (for example, muscle) which the damaged nerve innervates,
CC whereby the treatment of the muscle with MGF or a polynucleotide encoding
CC MGF prevents or diminishes degeneration. The method is useful for
CC treating neurological disorders, preferably motoneuron disorders. These
CC methods can reduce motoneuron loss by 20% or greater in response to nerve
CC avulsion. This sequence represents cDNA encoding the rabbit MGF.
XX
SQ Sequence 523 BP; 154 A; 129 C; 142 G; 98 T; 0 other;

```

```

Query Match          66.2%; Score 356.8; DB 24; Length 523;
Best Local Similarity 82.3%; Pred. No. 6.3e-91;
Matches 436; Conservative 0; Mismatches 87; Indels 7; Gaps 2;

```

```

Qy      1 GGACCAGAGACCCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTTCGTGTGTGGACCA 60
      ||||| ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1 GGACCGGAGACGCTCTGCGGTGCTGAGCTGGTGGATGCTCTTCAGTTTCGTGTGTGGAGAC 60

Qy     61 AGGGGCTTTTACTTCAACAAGCCACAGTCTATGGCTCCAGCATTTCGGAGGGGCACCACAG 120
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     61 AGGGGCTTTTATTTCACAACAAGCCACAGGATACGGCTCCAGCAGTCGGAGGGGCACCTCAG 120

Qy    121 ACGGGCATTGTGGATGAGTGTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 180
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    121 ACAGGCATCGTGGATGAGTGTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 180

```

```

Qy      181 TGTGTCCGCTGCAAGCCTACAAAGTCAGCTCGTTCCATCCGGGGCCAGCGCCCACTGAC 240
        |||| | | ||||| |||| | || | ||| ||||| ||||| |||||
Db      181 TGTGCACCCCTCAAGCCGGCAAAGGCAGCCCGCTCCGTCCGTGCCAGCGCCACACCGAC 240

Qy      241 ATGCCCAAGACTCAGAAGTCCCAGCCCCTATCGACACACAAGAAAAGGAAGCTGCAAAGG 300
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      241 ATGCCCAAGACTCAGAAGTATCAGCCTCCATCTACCAACAAGAAAATGAAGTCTCAGAGG 300

Qy      301 AGAAGGAAAGGAAGTACACTTGAAGAACACAAGTAGAGGAAGTGCAGGAAACAAGACCTA 360
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      301 AGAAGGAAAGGAAGTACATTTGAAGAACACAAGTAGAGGGAGTGCAGGAAACAAGAACTA 360

Qy      361 CAGAATGTAGGAGGAGCCTCCCGAGGAAACAGAAAATGCCACGTCACCGCAAGATCCTTTG 420
        ||| ||||| || || || | |||| | || | || | ||||| || |||||
Db      361 CAGGATGTAGGAAGACCCCTTCTGAGGAGTGAAGAAGGACAGGCCACCGCAGGACCCTTTG 420

Qy      421 CTGCTTGAGCAACCTGCAAAACATCGGAACACCTGCCAAATATCAATAATGAGTTCAATA 480
        || | ||||| |||| | || | |||| | || || | |||| | ||
Db      421 CTCTGCACAGTTACCTGTAAACATTGGAATACCGGCCA-----AAAAATAAGTTTGATC 474

Qy      481 TCATTTTCAGAGATGGGCATTTCCCTCAATGAAATACACAAGTAAACATTTC 530
        ||||| |||| | ||||| ||||| ||||| ||||| ||||| |||||
Db      475 ACATTTCAAAGAT-GGCATTTCCCCCAATGAAATACACAAGTAAACATTTC 523

```

RESULT 9

AAT84893

ID AAT84893 standard; cDNA; 553 BP.

XX

AC AAT84893;

XX

DT 14-APR-1998 (first entry)

XX

DE Rabbit insulin like growth factor 1 encoding cDNA.

XX

KW Insulin like growth factor 1; IGF-1; Ec peptide; muscle disorder;

KW heart; neuromuscular disease; primer; ss.

XX

OS Oryctolagus cuniculus.

XX

FH	Key	Location/Qualifiers
----	-----	---------------------

FT	CDS	1..366
----	-----	--------

FT		/*tag= a
----	--	----------

FT		/product= "IGF-1"
----	--	-------------------

XX

PN WO9733997-A1.

XX

PD 18-SEP-1997.

XX

PF 11-MAR-1997; 97WO-GB00658.

XX

PR 11-MAR-1996; 96GB-0005124.

XX

PA (UNLO ) ROYAL FREE HOSPITAL SCHOOL MED.

XX

PI Goldspink G;

```

XX
DR   WPI; 1997-470877/43.
DR   P-PSDB; AAW23301.
XX
PT   Use of insulin like growth factor I characterised by presence of Ec
PT   peptide - to treat humans or animals, particularly muscle disorders,
PT   heart conditions or neuromuscular diseases
XX
PS   Disclosure; Fig 3; 33pp; English.
XX
CC   A use of insulin like growth factor I (IGF-1) has been developed, and
CC   is characterised by the presence of the Ec peptide, or a functional
CC   equivalent, in the treatment or therapy of a human or animal. The IGF-1
CC   polypeptide can be used to treat muscular disorders, e.g. Duchenne or
CC   Becker muscular dystrophy, autosomal dystrophies and related progressive
CC   skeletal muscle weakness and wasting, muscle atrophy in ageing humans,
CC   spinal cord injury induced muscle atrophy and neuromuscular diseases,
CC   and cardiac disorders, e.g. diseases where promotion of cardiac muscle
CC   protein synthesis is a beneficial treatment, cardiomyopathies and acute
CC   heart failure or insult, specifically myocarditis or myocardial
CC   infarction. It can also be used to promote bone fracture healing and
CC   maintenance of bone in old age. The present sequence encodes rabbit
CC   IGF-1 used in the present specification.
XX
SQ   Sequence 553 BP; 159 A; 142 C; 147 G; 105 T; 0 other;

Query Match          66.2%; Score 356.8; DB 18; Length 553;
Best Local Similarity 82.3%; Pred. No. 6.4e-91;
Matches 436; Conservative 0; Mismatches 87; Indels 7; Gaps 2;

Qy      1 GGACCAGAGACCCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTTCGTGTGTGGACCA 60
      ||||| ||||| || ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      31 GGACCGGAGACGCTCTGCGGTGCTGAGCTGGTGGATGCTCTTCAGTTTCGTGTGTGGAGAC 90

Qy      61 AGGGGCTTTTACTTCAACAAGCCCACAGTCTATGGCTCCAGCATTCCGAGGGGCACACAG 120
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      91 AGGGGCTTTTATTTCACAAGCCCACAGGATACGGCTCCAGCAGTCCGAGGGGCACCTCAG 150

Qy     121 ACGGGCATTTGTGGATGAGTGTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 180
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     151 ACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 210

Qy     181 TGTGTCCGCTGCAAGCCTACAAAGTCAGCTCGTTCCATCCGGGCCCAGCGCCACACTGAC 240
      |||| | | ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     211 TGTGCACCCCTCAAGCCGGCAAAGGCAGCCGCTCCGTCCGTGCCCAGCGCCACACCGAC 270

Qy     241 ATGCCCAAGACTCAGAAGTCCCAGCCCCCTATCGACACACAAGAAAAGGAAGCTGCAAAGG 300
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     271 ATGCCCAAGACTCAGAAGTATCAGCCTCCATCTACCAACAAGAAAATGAAGTCTCAGAGG 330

Qy     301 AGAAGGAAAGGAAGTACACTTGAAGAACACAAGTAGAGGAAGTGCAGGAAACAAGACCTA 360
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     331 AGAAGGAAAGGAAGTACATTTGAAGAACACAAGTAGAGGGAGTGCAGGAAACAAGAACTA 390

Qy     361 CAGAATGTAGGAGGAGCCTCCCGAGGAACAGAAAATGCCACGTACCCGCAAGATCCTTTG 420
      ||| ||||| || ||||| ||||| ||||| ||||| ||||| |||||
Db     391 CAGGATGTAGGAAGACCCCTTCTGAGGAGTGAAGAAGGACAGGCCACCGCAGGACCCTTTG 450

```



PS Claim 4; Page 49-50; 66pp; English.

CC The present invention relates to use of mechano-growth factor (MGF),  
CC an Insulin-like Growth Factor-I (IGF-I) isoform in the manufacture of a  
CC medicament for the treatment of neurological disorder. The MGF is capable  
CC of reducing motoneurone loss by 20% or greater in response to nerve  
CC avulsion, and effects motoneurone rescue, preferably adult motoneurone  
CC rescue. The MGF polynucleotide and polypeptide are useful in the  
CC manufacture of a medicament for the treatment of a neurological disorder,  
CC including a disorder of motoneurons and/or neurodegenerative disorder,  
CC e.g., amyotrophic lateral sclerosis, spinal muscular atrophy, progressive  
CC spinal muscular atrophy, infantile or juvenile muscular atrophy,  
CC poliomyelitis or post-polio syndrome, a disorder caused by exposure to a  
CC toxin, motoneurone trauma, a motoneurone lesion or nerve damage, an  
CC injury that affects motoneurons, motoneurone loss associated with aging,  
CC autosomal or sex-linked muscular dystrophy, diabetic neuropathy,  
CC peripheral neuropathies, Alzheimer's disease and Parkinson's disease.  
CC The present sequence is human IGF-I isoform MGF cDNA. MGF is a muscle  
CC isoform having extracellular (Ec) domain, hence also referred as  
CC IGF-I-Ec. The MGF protein comprises amino acid sequences encoded by  
CC nucleic acid sequence of IGF-I exons 4, 5 and 6 in the reading frame  
CC of MGF.

SO Sequence 517 BP; 150 A; 130 C; 139 G; 98 T; 0 other:

Query Match 60.3%; Score 325.2; DB 22; Length 517;  
Best Local Similarity 81.2%; Pred. No. 5.4e-82;  
Matches 429; Conservative 0; Mismatches 88; Indels 11; Gaps 4;

Qy	1	GGACCAGAGACCCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTTCGTGTGTGGACCA	60
Db	1	GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTTCGTGTGTGGAGAC	60
Qy	61	AGGGGCTTTTTACTTCAACAAGCCCACAGTCTATGGCTCCAGCATTCCGAGGGGCACCACAG	120
Db	61	AGGGGCTTTTTATTTCACAACAAGCCCACAGGGTATGGCTCCAGCAGTCCGAGGGGCGCCTCAG	120
Qy	121	ACGGGCATTGTGGATGAGTGTTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC	180
Db	121	ACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT	180
Qy	181	TGTGTCCGCTGCAAGCCTACAAAGTCAGCTCGTTCCATCCGGGCCACGCCCACTGAC	240
Db	181	TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCACGCCACACCGAC	240
Qy	241	ATGCCCAAGACTCAGAAGTCCAGCCCCCTATCGACACACAAGAAAAGGAAGCTGCAAAGG	300
Db	241	ATGCCCAAGACCCAGAAGTATCAGCCCCCATCTACCAACAAGAACCGAAGTCTCA---G	297
Qy	301	AGAAGGAAAGGAAGTACACTTGAAGAACACAAGTAGAGGAAGTGCAGGAAAACAAGACCTA	360
Db	298	AGAAGGAAAGGAAGTACATTTGAAGAACACAAGTAGAGGGAGTGCAGGAAAACAAGAACTA	357
Qy	361	CAGAATGTAGGAGGAGCCTCCCGAGGAACAGAAAATGCCACGTCACCGCAAGATCCTTTG	420
Db	358	CAGGATGTA-GAAGACCCCTCTGAGGAGTGAAGAAGGACAGGCCACCGCAGGACCCCTTTG	416









SQ Sequence 671 BP; 160 A; 181 C; 180 G; 150 T; 0 other;

Query Match 55.6%; Score 299.6; DB 24; Length 671;  
Best Local Similarity 97.1%; Pred. No. 1e-74;  
Matches 305; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

```
Qy      1 GGACCAGAGACCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTCGTGTGTGGACCA 60
      |||
Db     140 GGACCAGAGACCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAATTCGTGTGTGGACCA 199

Qy      61 AGGGGCTTTTACTTCAACAAGCCACAGTCTATGGGCTCCAGCATTTCGGAGGGCACCACAG 120
      |||
Db     200 AGGGGCTTTTACTTCAACAAGCCACAGGCTATGGGCTCCAGCATTTCGGAGGGCACCACAG 259

Qy     121 ACGGGCATTGTGGATGAGTGTGTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 180
      |||
Db     260 ACGGGCATTGTGGATGAGTGTGTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 319

Qy     181 TGTGTCCGCTGCAAGCCTACAAAGTCAGCTCGTTCCATCCGGGCCCAGCGCCCACTGAC 240
      |||
Db     320 TGTGTCCGCTGCAAGCCTACAAAGTCAGCTCGTTCCATCCGGGCCCAGCGCCCACTGAC 379

Qy     241 ATGCCCAAGACTCAGAAGTCCCAGCCCCTATCGACACACAAGAAAAGGAAGCTGCAAAGG 300
      |||
Db     380 ATGCCCAAGACTCAGAAGTCCCAGCCCCTATCGACACACAAGAAAAGGAAGCTGCAAAGG 439

Qy     301 AGAAGGAAAGGAAG 314
      |||
Db     440 AGAAGGAAAGGAAG 453
```

RESULT 13

AAN70436

ID AAN70436 standard; cDNA; 818 BP.

XX

AC AAN70436;

XX

DT 25-MAR-2003 (updated)

DT 05-APR-1991 (first entry)

XX

DE Sequence encoding insulin-like growth factor 1A (IGF-1A).

XX

KW Growth promoter; lactation enhancer; cell proliferation; ss.

XX

OS Homo sapiens.

XX

PN EP229750-A.

XX

PD 22-JUL-1987.

XX

PF 06-JAN-1987; 87EP-0870001.

XX

PR 20-NOV-1986; 86US-0929671.

PR 07-JAN-1986; 86US-0816662.

XX

PA (UNIW ) UNIV WASHINGTON.

XX

PI Krivi GG, Rotwein PS;  
 XX  
 DR WPI; 1987-200203/29.  
 XX  
 PT New pre-pro-insulin-like growth factor-1 protein - obtd. by  
 PT recombinant DNA procedures for use as growth promoters for  
 PT enhancing lactation, for stimulating cell proliferation etc.  
 XX  
 PS Example; Fig 5; 59pp; English.  
 XX  
 CC A 42 base oligonucleotide corresponding to the DNA sequence encoding  
 CC amino acids 10 to 23 of mature human IGF-I was synthesized (AAN70437).  
 CC The radiolabeled 42 mer was then employed to screen for IGF-I  
 CC containing DNA sequences in a human liver cDNA library. Insulin-  
 CC like growth factors-1A and -1B cDNAs were isolated from a human cDNA  
 CC library by using lambda gt 11 (AAN70435, AAN70436). The human IGF-1  
 CC genomic gene was isolated and mapped. It encodes at least two  
 CC preproinsulin-like growth factor-1 proteins. An essentially pure  
 CC preproinsulin-like growth factor-1 protein comprising the sequence  
 CC of amino acids shown in Figure six is claimed (AAP70277).  
 CC (Updated on 25-MAR-2003 to correct PA field.)  
 XX  
 SQ Sequence 818 BP; 232 A; 186 C; 187 G; 213 T; 0 other;

Query Match 50.9%; Score 274.2; DB 8; Length 818;  
 Best Local Similarity 76.0%; Pred. No. 1.7e-67;  
 Matches 405; Conservative 0; Mismatches 68; Indels 60; Gaps 3;

```

Qy      1 GGACCAGAGACCCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTCGTGTGTGGACCA 60
      ||||| ||||| || ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      203 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 262

Qy      61 AGGGGCTTTTACTTCAACAAGCCACAGTCTATGGCTCCAGCATTCGGAGGGCACCACAG 120
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      263 AGGGGCTTTTATTTCAACAAGCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 322

Qy      121 ACGGGCATTGTGGATGAGTGTTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 180
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      323 ACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 382

Qy      181 TGTGTCCGCTGCAAGCCTACAAAGTCAGCTCGTTCCATCCGGGCCCAGCGCCCACTGAC 240
      || | | | ||||| ||||| ||||| || ||||| ||||| |||||
Db      383 TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCAGCGCCACACCGAC 442

Qy      241 ATGCCCAAGACTCAGAAGTCCCAGCCCCTATCGACACACAAGAAAAGGAAGCTGCAAAGG 300
      ||||| ||||| |||
Db      443 ATGCCCAAGACCCAG----- 457

Qy      301 AGAAGGAAAGGAAGTACACTTGAAGAACACAAGTAGAGGAAGTGCAGGAAACAAGACCTA 360
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      458 -----AAGGAAGTACATTTGAAGAACGCAAGTAGAGGGAGTGCAGGAAACAAGAACTA 510

Qy      361 CAGAATGTAGGAGGAGCCTCCCGAGGAACAGAAAATGCCACGTCACCGCAAGATCCTTTG 420
      ||| ||||| || ||||| ||||| || ||||| ||||| |||||
Db      511 CAGGATGTAGGAAGACCCCTCCTGAGGAGTGAAGAGTGACATGCCACCGCAGGATCCTTTG 570

Qy      421 --CTGCTTGAGCAACCTGCAAAACATCGGAACACCTGCCAAATATCAATAATGAGTTCAA 478
  
```



CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub.published\_pct\_sequences.

XX

SQ Sequence 7260 BP; 2330 A; 1415 C; 1240 G; 2275 T; 0 other;

Query Match 50.9%; Score 274.2; DB 24; Length 7260;  
Best Local Similarity 76.0%; Pred. No. 3.6e-67;  
Matches 405; Conservative 0; Mismatches 68; Indels 60; Gaps 3;

```
Qy      1 GGACCAGAGACCCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTCGTGTGTGGACCA 60
      ||||| ||||| || ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     311 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 370

Qy     61 AGGGGCTTTTACTTCAACAAGCCACAGTCTATGGCTCCAGCATTTCGGAGGGCACCACAG 120
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     371 AGGGGCTTTTATTTCAACAAGCCACAGGGTATGGCTCCAGCAGTTCGGAGGGCGCCTCAG 430

Qy    121 ACGGGCATTGTGGATGAGTGTGTGCTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 180
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     431 ACAGGCATCGTGGATGAGTGTGCTGCTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 490

Qy    181 TGTGTCCGCTGCAAGCCTACAAAGTCAGCTCGTTCCATCCGGGGCCAGCGCCCACTGAC 240
      || | | | ||||| ||||| ||||| || ||||| ||||| |||||
Db     491 TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCAGCGCCACACCGAC 550

Qy    241 ATGCCCAAGACTCAGAAGTCCCAGCCCCTATCGACACACAAGAAAAGGAAGCTGCAAAGG 300
      ||||| ||||| |||
Db     551 ATGCCCAAGACCCAG----- 565

Qy    301 AGAAGGAAAGGAAGTACACTTGAAGAACACAAGTAGAGGAAGTGCAGGAAACAAGACCTA 360
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     566 -----AAGGAAGTACATTTGAAGAACGCAAGTAGAGGGAGTGCAGGAAACAAGAACTA 618

Qy    361 CAGAATGTAGGAGGAGCCTCCCGAGGAACAGAAAATGCCACGTCACCGCAAGATCCTTTG 420
      ||| ||||| || ||||| ||||| | | || | ||||| |||||
Db     619 CAGGATGTAGGAAGACCCCTCCTGAGGAGTGAAGAGTGACATGCCACCGCAGGATCCTTTG 678

Qy    421 --CTGCTTGAGCAACCTGCAAAACATCGGAACACCTGCCAAATATCAATAATGAGTTCAA 478
      |||| ||| |||| |||| ||||| ||||| || ||||| |||
Db     679 CTCTGCACGAGTTACCTGTTAAACTTTGGAACACCTACCA-----AAAAATAAGTTTGA 732

Qy    479 TATCATTTTCAGAGATGGGCATTTCCCTCAATGAAATACACAAGTAAACATTCC 531
      || ||||| | ||||| ||||| ||||| ||||| |||||
Db     733 TAACATTTTAAAGATGGGCGTTTCCCCCAATGAAATACACAAGTAAACATTCC 785
```

RESULT 15

ABK84583

ID ABK84583 standard; cDNA; 7260 BP.

XX

AC ABK84583;

XX

DT 14-AUG-2002 (first entry)

XX

DE Human cDNA differentially expressed in granulocytic cells #1154.

XX

KW Human; ss; granulocytic cell; DNA chip; bacterial infection;  
 KW viral infection; parasitic infection; protozoal infection;  
 KW fungal infection; sterile inflammatory disease; psoriasis;  
 KW rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;  
 KW cardiac reperfusion injury; renal reperfusion injury; ARDS;  
 KW adult respiratory distress syndrome; inflammatory bowel disease;  
 KW Crohn's disease; ulcerative colitis; periodontal disease;  
 KW granulocyte activation; chronic inflammation; allergy.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200228999-A2.  
 XX  
 PD 11-APR-2002.  
 XX  
 PF 03-OCT-2001; 2001WO-US30821.  
 XX  
 PR 03-OCT-2000; 2000US-237189P.  
 XX  
 PA (GENE-) GENE LOGIC INC.  
 XX  
 PI Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;  
 XX  
 DR WPI; 2002-435328/46.  
 XX  
 PT Detecting granulocyte activation by detecting differential expression  
 PT of genes associated with granulocyte activation, which serves as  
 PT diagnostic markers that is useful for monitoring disease states and  
 PT drug toxicity -  
 XX  
 PS Claim 1; SEQ ID No 1154; 114pp; English.  
 XX  
 CC The invention relates to detecting (M1) granulocyte (GC) activation  
 CC (GCA), by detecting the level of expression of gene(s) (Gs) identified by  
 CC DNA chip analysis as given in the specification, and comparing  
 CC the expression level to an expression level in an unactivated  
 CC GC, where differential expression of Gs is indicative of GCA.  
 CC Also included are modulating (M2) GA by contacting GC with an agent  
 CC that alters the expression of at least one gene in Gs; (2) screening (M3)  
 CC for an agent capable of modulating GCA or an inflammation (especially  
 CC chronic) in a tissue, an allergic response in a subject, exposure of a  
 CC subject to a pathogen or sterile inflammatory disease using the  
 CC gene expression profile; (3) detecting (M4) an inflammation (especially  
 CC chronic) in a tissue, an allergic response in a subject, exposure of a  
 CC subject to a pathogen or sterile inflammatory disease, by detecting the  
 CC level of expression in a sample of the tissue of gene(s) from Gs, where  
 CC the level of expression of the gene is indicative of inflammation;  
 CC (4) treating (M5) an inflammation (especially chronic) or in a tissue,  
 CC an allergic response in a subject, exposure of a subject to a pathogen  
 CC or sterile inflammatory disease, by contacting a tissue having  
 CC inflammation with an agent that modulates the expression of gene(s)  
 CC from Gs in the tissue. M1 is useful for detecting GCA; M2 is useful for  
 CC modulating GA; M3 is useful for screening an agent capable of modulating  
 CC GCA preferably in an inflammation in a tissue; M4 is useful for  
 CC detecting an inflammation (especially chronic) in a tissue, an allergic  
 CC response in a subject, exposure of a subject to a pathogen or sterile  
 CC inflammatory disease (e.g. psoriasis, rheumatoid arthritis,



glomerulonephritis, asthma, thrombosis, cardiac reperfusion injury, renal  
reperfusion injury, ARDS, adult respiratory distress syndrome,  
inflammatory bowel disease, Crohn's disease, ulcerative colitis,  
periodontal disease; also bacterial infection, viral infection,  
parasitic infection, protozoal infection, fungal infection and M5 is  
useful for treating one of the above conditions. The present  
sequence represents a gene differentially expressed in granulocytes.  
Note: The sequence data for this patent did not form part  
of the printed specification, but was obtained in electronic  
format directly from WIPO at  
[ftp.wipo.int/pub/published\\_pct\\_sequences](ftp.wipo.int/pub/published_pct_sequences).

Sequence 7260 BP; 2330 A; 1415 C; 1240 G; 2275 T; 0 other:

Query Match 50.9%; Score 274.2; DB 24; Length 7260;  
Best Local Similarity 76.0%; Pred. No. 3.6e-67;  
Matches 405; Conservative 0; Mismatches 68; Indels 60; Gaps 3;

QY            1 GGACCAGAGACCCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTCGTGTGTGGACCA 60  
             | | | | | | | | | | | | | | | | | | | | | | | | | | | |

[illegible]

Qy 121 ACGGGCATTGTGGATGAGTGTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 180

Qy 181 TGTGTCCGCTGCAAGCCTACAAAGTCAGCTCGTTCCATCCGGGGCCAGCGCCACACTGAC 240

Qy 241 ATGCCCCAAGACTCAGAAGTCCCAGCCCCTATCGACACACAAGAAAAGGAAGCTGCAAAGG 300  
|||||

Qy 301 AGAAGGAAAGGAAGTACACTTGAAGAACAAGTAGAGGAAGTGCAGGAAACAAGACCTA 360

Qy 361 CAGAATGTAGGAGGAGCCTCCCGAGGAACAGAAAATGCCACGTACCCGCAAGATCCTTTG 420

Qy 421 --CTGCTTGAGCAACCTGCAAAACATCGGAACACCTGCCAAATATCAATAATGAGTTCAA 478

QY 479 TATCATTTT CAGAGATGGGCATTTCCCTCAATGAAATACACAAGTAAACATTCC 531

Search completed: December 13, 2003, 06:03:49

Job time : 217.419 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 13, 2003, 06:03:55 ; Search time 49.8379 Seconds  
(without alignments)  
4773.589 Million cell updates/sec

Title: US-09-852-261-3  
Perfect score: 539  
Sequence: 1 ggaccagagaccctttgcgg.....agtaaacattcccgaattc 539

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_NA:\*  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/PCTUS\_COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	%		DB	ID	Description
		Query	Match Length			
1	356.8	66.2	553	3	US-09-142-583A-3	Sequence 3, Appli
2	356.8	66.2	553	3	US-09-142-583A-5	Sequence 5, Appli
3	272.6	50.6	777	3	US-09-142-583A-10	Sequence 10, Appli
4	271	50.3	622	6	5405942-2	Patent No. 5405942
5	222	41.2	5707	2	US-08-472-809B-8	Sequence 8, Appli
6	222	41.2	6345	2	US-08-472-809B-7	Sequence 7, Appli
7	205.2	38.1	357	6	5405942-9	Patent No. 5405942
8	203.6	37.8	357	6	5405942-13	Patent No. 5405942
9	166.8	30.9	210	6	5405942-7	Patent No. 5405942
10	166.8	30.9	210	6	5405942-11	Patent No. 5405942
11	166.8	30.9	2862	4	US-09-255-829-13	Sequence 13, Appli

12	165.2	30.6	210	6	5405942-15	Patent No. 5405942
13	161.2	29.9	240	1	US-08-308-196A-1	Sequence 1, Appli
14	161.2	29.9	240	5	PCT-US91-06452-1	Sequence 1, Appli
15	161.2	29.9	390	3	US-09-029-267-13	Sequence 13, Appl
16	155.4	28.8	798	1	US-07-953-230A-6	Sequence 6, Appli
17	146	27.1	770	1	US-07-953-230A-1	Sequence 1, Appli
18	146	27.1	846	1	US-07-953-230A-5	Sequence 5, Appli
19	118.6	22.0	621	3	US-08-989-251-40	Sequence 40, Appl
20	118.6	22.0	621	3	US-09-340-250-40	Sequence 40, Appl
21	118.6	22.0	621	4	US-09-528-108-40	Sequence 40, Appl
22	110.2	20.4	243	2	US-08-482-182-75	Sequence 75, Appl
23	109.8	20.4	485	1	US-07-989-845-29	Sequence 29, Appl
24	109.8	20.4	485	1	US-07-989-844-13	Sequence 13, Appl
25	109.8	20.4	485	1	US-08-110-663-1	Sequence 1, Appli
26	109.8	20.4	485	1	US-08-169-688-1	Sequence 1, Appli
27	109.8	20.4	485	1	US-08-240-121-13	Sequence 13, Appl
28	109.8	20.4	485	1	US-08-451-241-13	Sequence 13, Appl
29	109.8	20.4	485	1	US-08-110-664-1	Sequence 1, Appli
30	109.8	20.4	485	1	US-08-446-882-1	Sequence 1, Appli
31	109.8	20.4	485	1	US-08-385-187A-1	Sequence 1, Appli
32	109.8	20.4	485	1	US-08-470-108-1	Sequence 1, Appli
33	109.8	20.4	485	5	PCT-US93-11297-13	Sequence 13, Appl
34	109.8	20.4	485	5	PCT-US93-11298-29	Sequence 29, Appl
35	109.2	20.3	237	1	US-07-764-655D-8	Sequence 8, Appli
c 36	107.6	20.0	237	1	US-07-764-655D-9	Sequence 9, Appli
37	107.6	20.0	717	1	US-08-284-784-40	Sequence 40, Appl
38	107.6	20.0	717	2	US-08-854-811-40	Sequence 40, Appl
39	107.6	20.0	783	1	US-08-284-784-43	Sequence 43, Appl
40	107.6	20.0	783	2	US-08-854-811-43	Sequence 43, Appl
41	107.6	20.0	891	1	US-08-284-784-33	Sequence 33, Appl
42	107.6	20.0	891	1	US-08-284-784-34	Sequence 34, Appl
43	107.6	20.0	891	2	US-08-854-811-33	Sequence 33, Appl
44	107.6	20.0	891	2	US-08-854-811-34	Sequence 34, Appl
45	107.6	20.0	900	1	US-08-284-784-24	Sequence 24, Appl

#### ALIGNMENTS

#### RESULT 1

US-09-142-583A-3

; Sequence 3, Application US/09142583A

; Patent No. 6221842

; GENERAL INFORMATION:

; APPLICANT: GOLDSPIK, GEOFFREY

; TITLE OF INVENTION: METHOD OF TREATING MUSCULAR DISORDERS

; NUMBER OF SEQUENCES: 11

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: NIXON & VANDERHYE P.C.

; STREET: 1100 NORTH GLEBE ROAD

; CITY: ARLINGTON

; STATE: VA

; COUNTRY: USA

; ZIP: 22201

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

```

;
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.25
;
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/142,583A
;   FILING DATE: 29-Oct-1998
;   CLASSIFICATION: <Unknown>
;
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: WO PCT/GB97/00658
;   FILING DATE: 11-MAR-1997
;   APPLICATION NUMBER: GB 9605124.8
;   FILING DATE: 11-MAR-1996
;
;   ATTORNEY/AGENT INFORMATION:
;   NAME: SADOFF, B. J.
;   REGISTRATION NUMBER: 36663
;   REFERENCE/DOCKET NUMBER: 117-263
;
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 7038164000
;   TELEFAX: 7038164100
;
;   INFORMATION FOR SEQ ID NO: 3:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 553 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: both
;   TOPOLOGY: linear
;   MOLECULE TYPE: cDNA
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 1..363
;   SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-142-583A-3

```

```

Query Match          66.2%;  Score 356.8;  DB 3;  Length 553;
Best Local Similarity 82.3%;  Pred. No. 2.5e-109;
Matches 436;  Conservative 0;  Mismatches 87;  Indels 7;  Gaps 2;

```

```

Qy      1 GGACCAGAGACCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTTCGTGTGTGGACCA 60
        ||||| ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      31 GGACCGGAGACGCTCTGCGGTGCTGAGCTGGTGGATGCTCTTCAGTTTCGTGTGTGGAGAC 90

Qy      61 AGGGGCTTTTACTTCAACAAGCCCACAGTCTATGGCTCCAGCATTCGGAGGGGCACCACAG 120
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      91 AGGGGCTTTTATTTCAACAAGCCCACAGGATACGGCTCCAGCAGTCGGAGGGGCACCTCAG 150

Qy     121 ACGGGCATTGTGGATGAGTGTTTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 180
        || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     151 ACAGGCATCGTGGATGAGTGCTTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 210

Qy     181 TGTGTCCGCTGCAAGCCTACAAAGTCAGCTCGTTCCATCCGGGGCCAGCGCCCACTGAC 240
        ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     211 TGTGCACCCCTCAAGCCGGCAAAGGCAGCCCGCTCCGTCCGTGCCAGCGCCACACCGAC 270

Qy     241 ATGCCCAAGACTCAGAAGTCCCAGCCCCTATCGACACACAAGAAAAGGAAGCTGCAAAGG 300
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     271 ATGCCCAAGACTCAGAAGTATCAGCCTCCATCTACCAACAAGAAAATGAAGTCTCAGAGG 330

Qy     301 AGAAGGAAAGGAAGTACACTTGAAGAACACAAGTAGAGGAAGTGCAGGAAACAAGACCTA 360
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```



```

; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 341..397
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-142-583A-5

```

```

Query Match          66.2%; Score 356.8; DB 3; Length 553;
Best Local Similarity 82.3%; Pred. No. 2.5e-109;
Matches 436; Conservative 0; Mismatches 87; Indels 7; Gaps 2;

```

```

Qy      1 GGACCAGAGACCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTTCGTGTGTGGACCA 60
      ||||| ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     31 GGACCGGAGACGCTCTGCGGTGCTGAGCTGGTGGATGCTCTTCAGTTTCGTGTGTGGAGAC 90

Qy     61 AGGGGCTTTTACTTCAACAAGCCACAGTCTATGGCTCCAGCATTTCGAGGGGCACCACAG 120
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     91 AGGGGCTTTTATTTCAACAAGCCACAGGATACGGCTCCAGCAGTTCGAGGGGCACCTCAG 150

Qy    121 ACGGGCATTGTGGATGAGTGTTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 180
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    151 ACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 210

Qy    181 TGTGTCCGCTGCAAGCCTACAAAGTCAGCTCGTTCCATCCGGGCCCAGCGCCACACTGAC 240
      ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    211 TGTGCACCCCTCAAGCCGGCAAAGGCAGCCGCTCCGTCCGTGCCCAGCGCCACACCGAC 270

Qy    241 ATGCCCAAGACTCAGAAGTCCCAGCCCTATCGACACACAAGAAAAGGAAGCTGCAAAGG 300
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    271 ATGCCCAAGACTCAGAAGTATCAGCCTCCATCTACCAACAAGAAAATGAAGTCTCAGAGG 330

Qy    301 AGAAGGAAAGGAAGTACACTTGAAGAACACAAGTAGAGGAAGTGCAGGAAACAAGACCTA 360
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    331 AGAAGGAAAGGAAGTACATTGAAGAACACAAGTAGAGGGAGTGCAGGAAACAAGAACTA 390

Qy    361 CAGAATGTAGGAGGAGCCTCCCGAGGAACAGAAAATGCCACGTCACCGCAAGATCCTTTG 420
      ||| ||||| || ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    391 CAGGATGTAGGAAGACCCTTCTGAGGAGTGAAGAAGGACAGGCCACCGCAGGACCCTTTG 450

Qy    421 CTGCTTGAGCAACCTGCAAAACATCGGAACACCTGCCAAATATCAATAATGAGTTCAATA 480
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    451 CTCTGCACAGTTACCTGTAAACATTGGAATACCGGCCA-----AAAAATAAGTTTGATC 504

Qy    481 TCATTTTCAGAGATGGGCATTTCCCTCAATGAAATACACAAGTAAACATTC 530
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    505 ACATTTCAAAGAT-GGCATTTCCCCCAATGAAATACACAAGTAAACATTC 553

```

# RESULT 3

US-09-142-583A-10

; Sequence 10, Application US/09142583A

; Patent No. 6221842

; GENERAL INFORMATION:

; APPLICANT: GOLDSPIK, GEOFFREY

; TITLE OF INVENTION: METHOD OF TREATING MUSCULAR DISORDERS

; NUMBER OF SEQUENCES: 11

; CORRESPONDENCE ADDRESS:

```

;      ADDRESSEE: NIXON & VANDERHYE P.C.
;      STREET: 1100 NORTH GLEBE ROAD
;      CITY: ARLINGTON
;      STATE: VA
;      COUNTRY: USA
;      ZIP: 22201
;
;      COMPUTER READABLE FORM:
;      MEDIUM TYPE: Floppy disk
;      COMPUTER: IBM PC compatible
;      OPERATING SYSTEM: PC-DOS/MS-DOS
;      SOFTWARE: PatentIn Release #1.0, Version #1.25
;
;      CURRENT APPLICATION DATA:
;      APPLICATION NUMBER: US/09/142,583A
;      FILING DATE: 29-Oct-1998
;      CLASSIFICATION: <Unknown>
;
;      PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: WO PCT/GB97/00658
;      FILING DATE: 11-MAR-1997
;      APPLICATION NUMBER: GB 9605124.8
;      FILING DATE: 11-MAR-1996
;
;      ATTORNEY/AGENT INFORMATION:
;      NAME: SADOFF, B. J.
;      REGISTRATION NUMBER: 36663
;      REFERENCE/DOCKET NUMBER: 117-263
;
;      TELECOMMUNICATION INFORMATION:
;      TELEPHONE: 7038164000
;      TELEFAX: 7038164100
;
;      INFORMATION FOR SEQ ID NO: 10:
;      SEQUENCE CHARACTERISTICS:
;      LENGTH: 777 base pairs
;      TYPE: nucleic acid
;      STRANDEDNESS: both
;      TOPOLOGY: linear
;      MOLECULE TYPE: cDNA
;      FEATURE:
;      NAME/KEY: CDS
;      LOCATION: 26..493
;      SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-09-142-583A-10.

```

```

Query Match          50.6%;  Score 272.6;  DB 3;  Length 777;
Best Local Similarity 75.8%;  Pred. No. 4.3e-81;
Matches 404;  Conservative 0;  Mismatches 69;  Indels 60;  Gaps 3;

```

```

Qy      1  GGACCAGAGACCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTCGTGTGTGGACCA 60
      ||||| ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      179 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 238

Qy      61  AGGGGCTTTTACTTCAACAAGCCACAGTCTATGGCTCCAGCATTCCGAGGGGCACCACAG 120
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      239 AGGGGCTTTTATTTCAACAAGCCACAGGGTATGGCTCCAGCAGTCCGAGGGGCGCCTCAG 298

Qy      121 ACGGGCATTGTGGATGAGTGTTTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 180
      || || || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      299 ACAGGTATCGTGGATGAGTGCTTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 358

Qy      181 TGTGTCCGCTGCAAGCCTACAAAGTCAGCTCGTTCCATCCGGGGCCAGCGCCCACTGAC 240

```







```

; FILING DATE: March 9, 1994
; APPLICATION NUMBER: 07/789,919
; FILING DATE: No. 5925564ember 6, 1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 214/212
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5707 bases
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-472-809B-8

```

```

Query Match          41.2%; Score 222; DB 2; Length 5707;
Best Local Similarity 75.5%; Pred. No. 1.1e-63;
Matches 314; Conservative 0; Mismatches 50; Indels 52; Gaps 1;

```

```

Qy      1 GGACCAGAGACCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTCGTGTGTGGACCA 60
      ||||| ||||| || ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      793 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 852

Qy      61 AGGGGCTTTTACTTCAACAAGCCACAGTCTATGGCTCCAGCATTCCGAGGGGCACACAG 120
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      853 AGGGGCTTTTATTTCAACAAGCCACAGGGTATGGCTCCAGCAGTCCGAGGGGCGCCTCAG 912

Qy      121 ACGGGCATTGTGGATGAGTGTGTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 180
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      913 ACAGGCATCGTGGATGAGTGTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 972

Qy      181 TGTGTCCGCTGCAAGCCTACAAAGTCAGCTCGTTCATCCGGGCCAGCGCCACACTGAC 240
      || | | | ||||| ||||| ||||| || || ||||| ||||| |||||
Db      973 TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGTCTGTCCGTGCCAGCGCCACACCGAC 1032

Qy      241 ATGCCCAAGACTCAGAAGTCCCAGCCCCTATCGACACACAAGAAAAGGAAGCTGCAAAGG 300
      ||||| ||||| |||
Db      1033 ATGCCCAAGACCCAG----- 1047

Qy      301 AGAAGGAAAGGAAGTACACTTGAAGAACACAAGTAGAGGAAGTGCAGGAAACAAGACCTA 360
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1048 -----AAGGAAGTACATTTGAAGAACGCAAGTAGAGGGAGTGCAGGAAACAAGAACTA 1100

Qy      361 CAGAATGTAGGAGGAGCCTCCCGAGGAACAGAAAATGCCACGTACCCGCAAGATCC 416
      ||| ||||| || ||||| ||||| | | | | | ||||| |||||
Db      1101 CAGGATGTAGGAAGACCCTCCTGAGGAGTGAAGAGTGACATGCCACCGCAGGATCC 1156

```

```

RESULT 6
US-08-472-809B-7
; Sequence 7, Application US/08472809B
; Patent No. 5925564

```

```

; GENERAL INFORMATION:
; APPLICANT: Schwartz, Robert J.
; APPLICANT: DeMayo, Franco J.
; APPLICANT: O'Malley, Bert W.
; TITLE OF INVENTION: Expression Vector Systems and
; TITLE OF INVENTION: Method of Use
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472,809B
; FILING DATE: June 7, 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/209,846
; FILING DATE: March 9, 1994
; APPLICATION NUMBER: 07/789,919
; FILING DATE: No. 5925564ember 6, 1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 214/212
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6345 bases
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-472-809B-7

```

```

Query Match          41.2%; Score 222; DB 2; Length 6345;
Best Local Similarity 75.5%; Pred. No. 1.2e-63;
Matches 314; Conservative 0; Mismatches 50; Indels 52; Gaps 1;

```

```

Qy      1 GGACCAGAGACCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTTCGTGTGTGGACCA 60
        ||||| ||||| || ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      3702 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTTCGTGTGTGGAGAC 3761
Qy      61 AGGGGCTTTTACTTCAACAAGCCCACAGTCTATGGCTCCAGCATTCGGAGGGGCACCACAG 120
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

Db      3762 AGGGGCTTTTATTTCAACAAGCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG 3821
QY      121 ACGGGCATTGTGGATGAGTGTTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 180
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      3822 ACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 3881
QY      181 TGTGTCCGCTGCAAGCCTACAAAGTCAGCTCGTTCCATCCGGGGCCAGCGCCACACTGAC 240
      || | | | ||||| | ||||| ||||| || ||| ||||| ||||| |||||
Db      3882 TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCAGCGCCACACCGAC 3941
QY      241 ATGCCCAAGACTCAGAAGTCCCAGCCCCTATCGACACACAAGAAAAGGAAGCTGCAAAGG 300
      ||||| ||||| |||
Db      3942 ATGCCCAAGACCCAG----- 3956
QY      301 AGAAGGAAAGGAAGTACACTTGAAGAACACAAGTAGAGGAAGTGCAGGAAACAAGACCTA 360
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      3957 -----AAGGAAGTACATTTGAAGAACGCAAGTAGAGGGAGTGCAGGAAACAAGAACTA 4009
QY      361 CAGAATGTAGGAGGAGCCTCCCGAGGAACAGAAAATGCCACGTCACCGCAAGATCC 416
      ||| ||||| || ||||| ||||| | | || || ||||| |||||
Db      4010 CAGGATGTAGGAAGACCTCCTGAGGAGTGAAGAGTGACATGCCACCGCAGGATCC 4065

```

# RESULT 7

5405942-9

;Patent No. 5405942

; APPLICANT: BELL, GRAEME I.; RALL, LESLIE B.; MERRYWEATHER,  
;JAMES P.

; TITLE OF INVENTION: PREPRO INSULIN-LIKE GROWTH FACTORS  
;I AND II

; NUMBER OF SEQUENCES: 16

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/65,673

; FILING DATE: 16-JUN-1987

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 630,557

; FILING DATE: 19-JUL-1984

;SEQ ID NO:9:

; LENGTH: 357

5405942-9

Query Match 38.1%; Score 205.2; DB 6; Length 357;

Best Local Similarity 70.2%; Pred. No. 9e-59;

Matches 181; Conservative 44; Mismatches 33; Indels 0; Gaps 0;

```

QY      1 GGACCAGAGACCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTTCGTGTGTGGACCA 60
      ||||| ||||| | : : ||||| : ||| : | : ||||| : : ||| : | : |||
Db      43 GGACCGGAGACGCUCUGCGGGGCGAGCUGGUGGACGCUCUUCAGUUCGUGUGUGGAGAC 102
QY      61 AGGGGCTTTTACTTCAACAAGCCACAGTCTATGGCTCCAGCATTTCGGAGGGCACCACAG 120
      ||||| : : : | : ||||| ||||| | : ||| : ||||| : ||||| || |||
Db      103 AGGGGCUUUUAUUUCAACAAGCCACAGGGUAUGGUCCAGCAGUCGAGGGCGCCUCAG 162
QY      121 ACGGGCATTGTGGATGAGTGTTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 180
      || || | : | : ||| : ||| : | : ||||| : ||| : | : ||||| : ||||| : |
Db      163 ACAGGUUAUCGUGGAUGAGUGCUGUUUCCGGAGCUGUGAUCUAAGGAGGCUGGAGAUGUAU 222

```

```

Qy      181 TGTGTCCGCTGCAAGCCTACAAAGTCAGCTCGTTCCATCCGGGGCCCAGCGCCACACTGAC 240
          :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      223 UGCGCACCCCUCAAGCCUGCCAAGUCAGCUCGCUCUGUCCGUGCCCAGCGCCACACCGAC 282

Qy      241 ATGCCCAAGACTCAGAAG 258
          |:| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      283 AUGCCCAAGACCCAGAAG 300

```

RESULT 8

5405942-13

;Patent No. 5405942

; APPLICANT: BELL, GRAEME I.; RALL, LESLIE B.; MERRYWEATHER,  
; JAMES P.

; TITLE OF INVENTION: PREPRO INSULIN-LIKE GROWTH FACTORS  
; I AND II

; NUMBER OF SEQUENCES: 16

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/65,673

; FILING DATE: 16-JUN-1987

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 630,557

; FILING DATE: 19-JUL-1984

;SEQ ID NO:13:

; LENGTH: 357

5405942-13

Query Match 37.8%; Score 203.6; DB 6; Length 357;

Best Local Similarity 86.8%; Pred. No. 3.1e-58;

Matches 224; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

```

Qy      1 GGACCAGAGACCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTCGTGTGTGGACCA 60
          ||||| ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      43 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 102

Qy      61 AGGGGCTTTTACTTCAACAAGCCCACAGTCTATGGCTCCAGCATTTCGGAGGGCACCACAG 120
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      103 AGGGGCTTTTATTTCAACAAGCCCACAGGGTATGGCTCCAGCAGTCGGAGGGGCGCCTCAG 162

Qy      121 ACGGGCATTGTGGATGAGTGTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 180
          || || || || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      163 ACAGGTATCGTGGATGAGTGTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 222

Qy      181 TGTGTCCGCTGCAAGCCTACAAAGTCAGCTCGTTCCATCCGGGGCCCAGCGCCACACTGAC 240
          || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      223 TGCGCACCCCTCAGGCCTGCCAAGTCAGCTCGTCTGTCCGTGCCAGCGCCACACCGAC 282

Qy      241 ATGCCCAAGACTCAGAAG 258
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      283 ATGCCCAAGACCCAGAAG 300

```

RESULT 9

5405942-7

;Patent No. 5405942

; APPLICANT: BELL, GRAEME I.; RALL, LESLIE B.; MERRYWEATHER,  
; JAMES P.

```
; TITLE OF INVENTION: PREPRO INSULIN-LIKE GROWTH FACTORS
; I AND II
; NUMBER OF SEQUENCES: 16
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/65,673
; FILING DATE: 16-JUN-1987
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 630,557
; FILING DATE: 19-JUL-1984
; SEQ ID NO:7:
; LENGTH: 210
5405942-7
```

```
Query Match          30.9%; Score 166.8; DB 6; Length 210;
Best Local Similarity 67.6%; Pred. No. 4.5e-46;
Matches 142; Conservative 41; Mismatches 27; Indels 0; Gaps 0;
```

```
Qy      1 GGACCAGAGACCCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTTCGTGTGTGGACCA 60
        ||||| ||||| |: :||| |||: |||: ||| |||: |||: |||: |||: |||
Db      1 GGACCGGAGACGCUCUGCGGGGUGAGCUGGUGGAUGCUCUUCAGUUCGUGUGUGGAGAC 60

Qy     61 AGGGGCTTTTACTTCAACAAGCCACAGTCTATGGCTCCAGCATTCGGAGGGCACCACAG 120
        |||||: :|||: |||: |||: |||: |||: |||: |||: |||: |||
Db     61 AGGGGCUUUUAUUUCAACAAGCCACAGGGUAUGGCUCACAGUCGAGGGGCGCCUCAG 120

Qy    121 ACGGGCATTGTGGATGAGTGTGTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 180
        || || |: |: |||: |||: |: |: |||: |||: |: |: |||: |||: |||: |: |:
Db    121 ACAGGUAUCGUGGAUGAGUGCUGCUCCGGAGCUGUGAUCUAAGGAGGCUGGAGAUGUAU 180

Qy    181 TGTGTCCGCTGCAAGCCTACAAAGTCAGCT 210
        :| | | | |||||: | |||: ||||:
Db    181 UGCGCACCCCUCAAGCCUGCCAAGUCAGCU 210
```

```
RESULT 10
5405942-11
; Patent No. 5405942
; APPLICANT: BELL, GRAEME I.; RALL, LESLIE B.; MERRYWEATHER,
; JAMES P.
; TITLE OF INVENTION: PREPRO INSULIN-LIKE GROWTH FACTORS
; I AND II
; NUMBER OF SEQUENCES: 16
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/65,673
; FILING DATE: 16-JUN-1987
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 630,557
; FILING DATE: 19-JUL-1984
; SEQ ID NO:11:
; LENGTH: 210
5405942-11
```

```
Query Match          30.9%; Score 166.8; DB 6; Length 210;
Best Local Similarity 87.1%; Pred. No. 4.5e-46;
Matches 183; Conservative 0; Mismatches 27; Indels 0; Gaps 0;
```

```
Qy      1 GGACCAGAGACCCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTTCGTGTGTGGACCA 60
```



```

;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;      MOLECULE TYPE:  DNA (genomic)
;      FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1..2862
US-09-255-829-13

```

Query Match 30.9%; Score 166.8; DB 4; Length 2862;  
Best Local Similarity 87.1%; Pred. No. 2.1e-45;  
Matches 183; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

Qy	1	GGACCAGAGACCCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTCGTGTGTGGACCA	60
Db	2644	GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC	2703
Qy	61	AGGGGCTTTTTACTTCAACAAGCCCACAGTCTATGGCTCCAGCATTCGGAGGGGCACCAACAG	120
Db	2704	AGGGGCTTTTTATTTCACAACAAGCCCACAGGGTATGGCTCCAGCAGTCGGAGGGGCGCCTCAG	2763
Qy	121	ACGGGCATTGTGGATGAGTGTTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC	180
Db	2764	ACAGGTATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT	2823
Qy	181	TGTGTCCGCTGCAAGCCTACAAAGTCAGCT	210
Db	2824	TGCGCACCCCTCAAGCCTGCCAAGTCAGCT	2853

RESULT 12  
5405942-15  
;Patent No. 5405942  
; APPLICANT: BELL, GRAEME I.;RALL, LESLIE B.;MERRYWEATHER,  
;JAMES P.  
; TITLE OF INVENTION: PREPRO INSULIN-LIKE GROWTH FACTORS  
;I AND II  
; NUMBER OF SEQUENCES: 16  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/65,673  
; FILING DATE: 16-JUN-1987  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 630,557  
; FILING DATE: 19-JUL-1984  
;SEQ ID NO:15:  
; LENGTH: 210  
5405942-15

Query Match 30.6%; Score 165.2; DB 6; Length 210;  
Best Local Similarity 67.1%; Pred. No. 1.5e-45;  
Matches 141; Conservative 41; Mismatches 28; Indels 0; Gaps 0;

Qy 1 GGACCAGAGACCCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTCGTGTGTGGACCA 60  
 ||||| ||||| |: :|||||||:||||:|:| | |:|:| ||:|:| ||  
 Db 1 GGACCGGAGACGCUCUGCGGGGCGAGCUGGUGGAUGCUCUUCAGUUCGUGUGUGGAGAC 60  
 Qy 61 AGGGGCTTTTACTTCAACAAGCCACAGTCTATGGCTCCAGCATTTCGAGGGGCACCACAG 120





US-08-308-196A-1

Matches 181; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

Qy 1 GGACCAGAGACCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTCGTGTGTGGACCA 60

## RESULT 14

PCT-US91-06452-1

; Sequence 1, Application PC/TUS9106452

GENERAL INFORMATION:

APPLICANT: Brierley, Russell A.

APPLICANT: Davis, Geneva R.

; APPLICANT: Holtz, Gregory C.

APPLICANT: Gleeson, Martin A.

APPLICANT: Bradley, D. H.

TITLE OF INVENTION: Production of Insulin-Like Growth

TITLE OF INVENTION: Factor-1 in Methyloleophilic Yeast Cells

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fitch, Even, Tabin & Flannery

STREET: 135 South LaSalle Street, Suite 900

; CITY: Chicago

; STATE: Illinois

; COUNTRY: USA

; ZIP: 60603

COMPUTER READABLE FORM:

```

; MEDIUM TYPE: Floppy disk

```

```

;      COMPUTER:  IBM PC compatible

```

OPERATING SYSTEM: PC-DOS/MS-DOS

```

; SOFTWARE: PatentIn Release #1.0, Version #1.25

```

; CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US91/06452

; FILING DATE: 19910409

CLASSIFICATION: 435

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/578,728
; FILING DATE: 04-SEP-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie L.
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 51874
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619)552-1311
; TELEFAX: (619)552-0095
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 240 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 14..232
PCT-US91-06452-1

```

```

Query Match          29.9%; Score 161.2; DB 5; Length 240;
Best Local Similarity 84.6%; Pred. No. 3.6e-44;
Matches 181; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

```

```

Qy      1 GGACCAGAGACCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTCGTGTGTGGACCA 60
      ||||| ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      17 GGACCGGAGACGCTCTGCGGGGCTGAGCTCGTGGATGCTCTGCAGTTCGTGTGTGGAGAC 76

Qy      61 AGGGGCTTTTACTTCAACAAGCCACAGTCTATGGCTCCAGCATTCGGAGGGCACCACAG 120
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      77 AGGGGCTTTTATTTCAACAAGCCACAGGGTATGGCTCCAGCAGTCGACGGGCGCCTCAG 136

Qy      121 ACGGGCATTGTGGATGAGTGTGTGCTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 180
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      137 ACAGGCATCGTGGATGAGTGTGCTTCCGGAGCTGTGATCTAAGGAGGCTCGAGATGTAT 196

Qy      181 TGTGTCCGCTGCAAGCCTACAAAGTCAGCTCGTT 214
      || | | | ||||| ||||| ||||| || |
Db      197 TGCGCACCCCTCAAGCCTGCCAAGTCAGCTTGAT 230

```

```

RESULT 15
US-09-029-267-13

```

```

; Sequence 13, Application US/09029267
; Patent No. 6107057
; GENERAL INFORMATION:
; APPLICANT: Crawford, Kenneth
; APPLICANT: Zaror, Isabel
; APPLICANT: Innis, Michael
; TITLE OF INVENTION: Pichia Secretory Leader for Protein
; TITLE OF INVENTION: Expression
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street

```

```

; CITY: Emeryville
; STATE: California
; COUNTRY: United States
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/029,267
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Chung, Ling-Fong
; REGISTRATION NUMBER: 36,482
; REFERENCE/DOCKET NUMBER: 1165.100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 601-2704
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 390 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "Synthetic"
US-09-029-267-13

```

```

Query Match          29.9%; Score 161.2; DB 3; Length 390;
Best Local Similarity 84.6%; Pred. No. 4.8e-44;
Matches 181; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

```

```

Qy      1 GGACCAGAGACCCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTCGTGTGTGGACCA 60
        ||||| ||||| || ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      160 GGACCGGAGACGCTCTGCGGGGCTGAGCTCGTGGATGCTCTGCAGTTCGTGTGTGGAGAC 219

Qy      61 AGGGGCTTTTACTTCAACAAGCCCAAGTCTATGGCTCCAGCATTCCGAGGGGCACACAG 120
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      220 AGGGGCTTTTATTTCAACAAGCCCAAGGGTATGGCTCCAGCAGTCCGAGGGCGCCTCAG 279

Qy      121 ACGGGCATTGTGGATGAGTGTGTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 180
        || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      280 ACAGGCATCGTGGATGAGTGTGCTTCCGGAGCTGTGATCTAAGGAGGCTCGAGATGTAT 339

Qy      181 TGTGTCCGCTGCAAGCCTACAAAGTCAGCTCGTT 214
        || | | | ||||| | ||||| |||||
Db      340 TGCGCACCCCTCAAGCCTGCCAAGTCAGCTTGAT 373

```

```

Search completed: December 13, 2003, 11:44:49
Job time : 49.8379 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 13, 2003, 07:29:55 ; Search time 240.655 Seconds  
(without alignments)  
7443.919 Million cell updates/sec

Title: US-09-852-261-3  
Perfect score: 539  
Sequence: 1 ggaccagagaccctttgcgg.....agtaaacattcccgaattc 539

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2201672 seqs, 1661799599 residues

Total number of hits satisfying chosen parameters: 4403344

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published\_Applications\_NA:\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq2:\*
- 14: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:\*
- 16: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*
- 17: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*
- 18: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	Score	Match	Length	ID	Description
No.					

1	539	100.0	539	9	US-09-852-261-3	Sequence 3, Appli
2	453.8	84.2	651	15	US-10-161-088-1	Sequence 1, Appli
3	425	78.8	487	9	US-09-852-261-11	Sequence 11, Appli
4	356.8	66.2	523	9	US-09-852-261-5	Sequence 5, Appli
5	325.2	60.3	517	9	US-09-852-261-1	Sequence 1, Appli
6	274.2	50.9	7260	10	US-09-919-497-24	Sequence 24, Appli
7	274.2	50.9	7260	10	US-09-880-107-3739	Sequence 3739, Ap
8	274.2	50.9	7260	13	US-09-873-319-707	Sequence 707, App
9	274.2	50.9	7260	13	US-09-960-706-1066	Sequence 1066, Ap
10	274.2	50.9	7260	15	US-10-136-639-4	Sequence 4, Appli
11	272.6	50.6	725	15	US-10-207-655-54	Sequence 54, Appli
12	262	48.6	471	9	US-09-852-261-13	Sequence 13, Appli
13	224	41.6	612	13	US-10-251-661-7	Sequence 7, Appli
14	213.8	39.7	286	15	US-10-161-088-3	Sequence 3, Appli
15	206.8	38.4	318	9	US-09-852-261-9	Sequence 9, Appli
16	203.6	37.8	462	15	US-10-238-114-1	Sequence 1, Appli
17	168.4	31.2	210	13	US-09-807-742-18	Sequence 18, Appli
18	166.8	30.9	2862	13	US-10-241-596-13	Sequence 13, Appli
19	164.6	30.5	4532	10	US-09-930-377B-1	Sequence 1, Appli
20	163.6	30.4	210	10	US-09-930-377B-2	Sequence 2, Appli
21	161.2	29.9	390	15	US-10-179-046-13	Sequence 13, Appli
22	143.6	26.6	182	13	US-10-029-386-18231	Sequence 18231, A
23	143.6	26.6	516	13	US-10-029-386-5832	Sequence 5832, Ap
24	124.2	23.0	213	15	US-10-076-816-9	Sequence 9, Appli
25	124.2	23.0	213	15	US-10-077-381-9	Sequence 9, Appli
26	118.6	22.0	621	9	US-09-921-398-40	Sequence 40, Appli
27	118.6	22.0	621	15	US-10-280-826-40	Sequence 40, Appli
28	105.4	19.6	480	9	US-09-921-398-38	Sequence 38, Appli
29	105.4	19.6	480	15	US-10-280-826-38	Sequence 38, Appli
30	91.6	17.0	210	13	US-09-807-742-19	Sequence 19, Appli
31	67.4	12.5	411	10	US-09-960-352-2082	Sequence 2082, Ap
c 32	67.2	12.5	437	15	US-10-066-543-663	Sequence 663, App
c 33	67.2	12.5	493	15	US-10-066-543-997	Sequence 997, App
c 34	67.2	12.5	518	15	US-10-066-543-1040	Sequence 1040, Ap
c 35	67.2	12.5	536	15	US-10-066-543-428	Sequence 428, App
36	67.2	12.5	543	15	US-10-136-841-1	Sequence 1, Appli
c 37	67.2	12.5	549	15	US-10-066-543-478	Sequence 478, App
c 38	67.2	12.5	574	9	US-09-922-217-918	Sequence 918, App
c 39	67.2	12.5	574	10	US-09-833-263-918	Sequence 918, App
c 40	67.2	12.5	574	14	US-10-025-380-918	Sequence 918, App
c 41	67.2	12.5	577	15	US-10-066-543-1137	Sequence 1137, Ap
c 42	67.2	12.5	579	15	US-10-066-543-1094	Sequence 1094, Ap
c 43	67.2	12.5	586	15	US-10-066-543-808	Sequence 808, App
44	67.2	12.5	1045	13	US-09-873-319-423	Sequence 423, App
45	67.2	12.5	1045	13	US-09-960-706-675	Sequence 675, App

#### ALIGNMENTS

#### RESULT 1

US-09-852-261-3

; Sequence 3, Application US/09852261

; Patent No. US20020083477A1

; GENERAL INFORMATION:

; APPLICANT: GOLDSPIK, GEOFFREY

; APPLICANT: TERENCE, GIORGIO  
 ; TITLE OF INVENTION: REPAIR OF NERVE DAMAGE  
 ; FILE REFERENCE: 117-351  
 ; CURRENT APPLICATION NUMBER: US/09/852,261  
 ; CURRENT FILING DATE: 2001-05-10  
 ; PRIOR APPLICATION NUMBER: GB 0011278.9  
 ; PRIOR FILING DATE: 2000-05-10  
 ; NUMBER OF SEQ ID NOS: 14  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 3  
 ; LENGTH: 539  
 ; TYPE: DNA  
 ; ORGANISM: Rattus sp.  
 US-09-852-261-3

Query Match 100.0%; Score 539; DB 9; Length 539;  
 Best Local Similarity 100.0%; Pred. No. 8.4e-169;  
 Matches 539; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	GGACCAGAGACCCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTTCGTGTGTGGACCA	60
Db	1	GGACCAGAGACCCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTTCGTGTGTGGACCA	60
Qy	61	AGGGGCTTTTACTTCAACAAGCCACAGTCTATGGCTCCAGCATTCGGAGGGCACCACAG	120
Db	61	AGGGGCTTTTACTTCAACAAGCCACAGTCTATGGCTCCAGCATTCGGAGGGCACCACAG	120
Qy	121	ACGGGCATTGTGGATGAGTGTTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC	180
Db	121	ACGGGCATTGTGGATGAGTGTTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC	180
Qy	181	TGTGTCCGCTGCAAGCCTACAAAGTCAGCTCGTTCCATCCGGGCCAGCGCCCACTGAC	240
Db	181	TGTGTCCGCTGCAAGCCTACAAAGTCAGCTCGTTCCATCCGGGCCAGCGCCCACTGAC	240
Qy	241	ATGCCCAAGACTCAGAAGTCCCAGCCCCTATCGACACACAAGAAAAGGAAGCTGCAAAGG	300
Db	241	ATGCCCAAGACTCAGAAGTCCCAGCCCCTATCGACACACAAGAAAAGGAAGCTGCAAAGG	300
Qy	301	AGAAGGAAAGGAAGTACACTTGAAGAACACAAGTAGAGGAAGTGCAGGAAACAAGACCTA	360
Db	301	AGAAGGAAAGGAAGTACACTTGAAGAACACAAGTAGAGGAAGTGCAGGAAACAAGACCTA	360
Qy	361	CAGAATGTAGGAGGAGCCTCCCGAGGAACAGAAAATGCCACGTACCGCAAGATCCTTTG	420
Db	361	CAGAATGTAGGAGGAGCCTCCCGAGGAACAGAAAATGCCACGTACCGCAAGATCCTTTG	420
Qy	421	CTGCTTGAGCAACCTGCAAAACATCGGAACACCTGCCAAATATCAATAATGAGTTCAATA	480
Db	421	CTGCTTGAGCAACCTGCAAAACATCGGAACACCTGCCAAATATCAATAATGAGTTCAATA	480
Qy	481	TCATTTTCAAGATGGGCATTTCCCTCAATGAAATACACAAGTAAACATTCCCGGAATTC	539
Db	481	TCATTTTCAAGATGGGCATTTCCCTCAATGAAATACACAAGTAAACATTCCCGGAATTC	539

US-10-161-088-1  
; Sequence 1, Application US/10161088  
; Publication No. US20030077761A1  
; GENERAL INFORMATION:  
; APPLICANT: Parrow, Vendela  
; APPLICANT: Rosengren, Linda  
; TITLE OF INVENTION: NEW METHODS  
; FILE REFERENCE: 13425-111001  
; CURRENT APPLICATION NUMBER: US/10/161,088  
; CURRENT FILING DATE: 2002-05-31  
; PRIOR APPLICATION NUMBER: SE 0101934-8  
; PRIOR FILING DATE: 2001-06-01  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 651  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (73)...(471)  
US-10-161-088-1

Query Match 84.2%; Score 453.8; DB 15; Length 651;  
Best Local Similarity 92.8%; Pred. No. 2.3e-140;  
Matches 476; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

```

Qy      1 GGACCAGAGACCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTTCGTGTGTGGACCA 60
      |||
Db      139 GGACCAGAGACCCTTTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTTCGTGTGTGGACCG 198

Qy      61 AGGGGCTTTTACTTCAACAAGCCACAGTCTATGGCTCCAGCATTTCGGAGGGCACCACAG 120
      |||
Db      199 AGGGGCTTTTACTTCAACAAGCCACAGGCTATGGCTCCAGCATTTCGGAGGGCACCCTCAG 258

Qy      121 ACGGGCATTGTGGATGAGTGTGTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 180
      ||
Db      259 ACAGGCATTGTGGATGAGTGTGTGCTTCCGGAGCTGTGATCTGAGGAGACTGGAGATGTAC 318

Qy      181 TGTGTCCGCTGCAAGCCTACAAAGTCAGCTCGTTCCATCCGGGCCCAGCGCCCACTGAC 240
      ||||
Db      319 TGTGCCCCACTGAAGCCTACAAAAGCAGCCCGCTCTATCCGTGCCCAGCGCCCACTGAC 378

Qy      241 ATGCCCAAGACTCAGAAGTCCCAGCCCCTATCGACACACAAGAAAAGGAAGCTGCAAAGG 300
      |||
Db      379 ATGCCCAAGACTCAGAAGTCCCCGTCCCTATCGACAAAACAAGAAAACGAAGCTGCAAAGG 438

Qy      301 AGAAGGAAAGGAAGTACACTTGAAGAACACAAGTAGAGGAAGTGCAAGAAACAAGACCTA 360
      |||
Db      439 AGAAGGAAAGGAAGTACATTTGAAGAACACAAGTAGAGGAAGTGCAAGAAACAAGACCTA 498

Qy      361 CAGAATGTAGGAGGAGCCTCCCGAGGAACAGAAAATGCCACGTCACCGCAAGATCCTTTG 420
      |||
Db      499 CAGAATGTAGGAGGAGCCTCCCGAGGAGCAGAAAATGCCACATCACCGCAGGATCCTTTG 558

Qy      421 CTGCTTGAGCAACCTGCAAAACATCGGAACACCTGCCAAATATCAATAATGAGTTCAATA 480
      |||

```



Db 559 CTGCTTGAGCAACCTGCAAAACATCGAAACACCTACCAAATAACAATAATAAGTCCAATA 618

Qy 481 TCATTTTCAGAGATGGGCATTTCCCTCAATGAAA 513

Db 619 ACATTACAAAGATGGGCATTTCCCCCAATGAAA 651

# RESULT 3

US-09-852-261-11

; Sequence 11, Application US/09852261

; Patent No. US20020083477A1

; GENERAL INFORMATION:

; APPLICANT: GOLDSPINK, GEOFFREY

; APPLICANT: TERENGHI, GIORGIO

; TITLE OF INVENTION: REPAIR OF NERVE DAMAGE

; FILE REFERENCE: 117-351

; CURRENT APPLICATION NUMBER: US/09/852,261

; CURRENT FILING DATE: 2001-05-10

; PRIOR APPLICATION NUMBER: GB 0011278.9

; PRIOR FILING DATE: 2000-05-10

; NUMBER OF SEQ ID NOS: 14

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 11

; LENGTH: 487

; TYPE: DNA

; ORGANISM: Rattus sp.

US-09-852-261-11

Query Match 78.8%; Score 425; DB 9; Length 487;

Best Local Similarity 90.4%; Pred. No. 7.6e-131;

Matches 487; Conservative 0; Mismatches 0; Indels 52; Gaps 1;

Qy 1 GGACCAGAGACCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTCGTGTGTGGACCA 60

Db 1 GGACCAGAGACCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTCGTGTGTGGACCA 60

Qy 61 AGGGGCTTTTACTTCAACAAGCCCACAGTCTATGGCTCCAGCATTCCGAGGGGCACCACAG 120

Db 61 AGGGGCTTTTACTTCAACAAGCCCACAGTCTATGGCTCCAGCATTCCGAGGGGCACCACAG 120

Qy 121 ACGGGCATTGTGGATGAGTGTGCTTCCGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 180

Db 121 ACGGGCATTGTGGATGAGTGTGCTTCCGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 180

Qy 181 TGTGTCCGCTGCAAGCCTACAAAGTCAGCTCGTTCCATCCGGGGCCAGCGCCCACTGAC 240

Db 181 TGTGTCCGCTGCAAGCCTACAAAGTCAGCTCGTTCCATCCGGGGCCAGCGCCCACTGAC 240

Qy 241 ATGCCCCAAGACTCAGAAGTCCCAGCCCCCTATCGACACACAAGAAAAGGAAGCTGCAAAGG 300

Db 241 ATGCCCCAAGACTCAG----- 255

Qy 301 AGAAGGAAAGGAAGTACACTTGAAGAACACAAGTAGAGGAAGTGCAGGAAACAAGACCTA 360

Db 256 -----AAGGAAGTACACTTGAAGAACACAAGTAGAGGAAGTGCAGGAAACAAGACCTA 308

Qy 361 CAGAATGTAGGAGGAGCCTCCCGAGGAACAGAAAATGCCACGTCACCGCAAGATCCTTTG 420

```

Db      309  |||||CAGAATGTAGGAGGAGCCTCCCGAGGAACAGAAAATGCCACGTCACCGCAAGATCCTTTG 368
Qy      421  CTGCTTGAGCAACCTGCAAAACATCGGAACACCTGCCAAATATCAATAATGAGTTCAATA 480
Db      369  |||||CTGCTTGAGCAACCTGCAAAACATCGGAACACCTGCCAAATATCAATAATGAGTTCAATA 428
Qy      481  TCATTTTCAGAGATGGGCATTTCCCTCAATGAAATACACAAGTAAACATTCCCGGAATTC 539
Db      429  |||||TCATTTTCAGAGATGGGCATTTCCCTCAATGAAATACACAAGTAAACATTCCCGGAATTC 487

```

RESULT 4

```

US-09-852-261-5
; Sequence 5, Application US/09852261
; Patent No. US20020083477A1
; GENERAL INFORMATION:
; APPLICANT: GOLDSPIK, GEOFFREY
; APPLICANT: TERENCE, GIORGIO
; TITLE OF INVENTION: REPAIR OF NERVE DAMAGE
; FILE REFERENCE: 117-351
; CURRENT APPLICATION NUMBER: US/09/852,261
; CURRENT FILING DATE: 2001-05-10
; PRIOR APPLICATION NUMBER: GB 0011278.9
; PRIOR FILING DATE: 2000-05-10
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 523
; TYPE: DNA
; ORGANISM: Oryctolagus cuniculus
US-09-852-261-5

```

```

Query Match          66.2%; Score 356.8; DB 9; Length 523;
Best Local Similarity 82.3%; Pred. No. 4.2e-108;
Matches 436; Conservative 0; Mismatches 87; Indels 7; Gaps 2;

```

```

Qy      1  GGACCAGAGACCCTTTGCGGGCTGAGCTGGTGGACGCTCTTCAGTTCGTGTGTGGACCA 60
Db      1  GGACCGGAGACGCTCTGCGGTGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 60
Qy     61  AGGGGCTTTTACTTCAACAAGCCACAGTCTATGGCTCCAGCATTCGGAGGGCACACAG 120
Db     61  AGGGGCTTTTATTTCAACAAGCCACAGGATACGGCTCCAGCAGTCGGAGGGCACCTCAG 120
Qy    121  ACGGGCATTGTGGATGAGTGTTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 180
Db    121  ACAGGCATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 180
Qy    181  TGTGTCCGCTGCAAGCCTACAAAGTCAGCTCGTTCCATCCGGGGCCAGCGCCACACTGAC 240
Db    181  TGTGCACCCCTCAAGCCGGCAAAGGCAGCCCGCTCCGTCCGTGCCAGCGCCACACCGAC 240
Qy    241  ATGCCCAAGACTCAGAAGTCCCAGCCCCTATCGACACACAAGAAAAGGAAGCTGCAAAGG 300
Db    241  ATGCCCAAGACTCAGAAGTATCAGCCTCCATCTACCAACAAGAAAATGAAGTCTCAGAGG 300

```

```

Qy      301 AGAAGGAAAGGAAGTACACTTGAAGAACAAGTAGAGGAAGTGCAGGAAACAAGACCTA 360
        |||||||||||||||| |||||||||||||||| |||||||||||||||| |||
Db      301 AGAAGGAAAGGAAGTACATTTGAAGAACAAGTAGAGGGAGTGCAGGAAACAAGAACTA 360

Qy      361 CAGAATGTAGGAGGAGCCTCCCGAGGAAACAGAAAATGCCACGTACCCGCAAGATCCTTTG 420
        ||| |||||||| || ||| ||||| || ||| ||| |||||||| || |||||||
Db      361 CAGGATGTAGGAAGACCCTTCTGAGGAGTGAAGAAGGACAGGCCACCGCAGGACCCTTTG 420

Qy      421 CTGCTTGAGCAACCTGCAAAACATCGGAACACCTGCCAAATATCAATAATGAGTTCAATA 480
        ||      | |||||| |||| |||| || |||| || |||| || ||
Db      421 CTCTGCACAGTTACCTGTAAACATTGGAATACCGGCCA-----AAAAATAAGTTTGATC 474

Qy      481 TCATTTTCAGAGATGGGCATTTCCCTCAATGAAATACACAAGTAAACATTC 530
        |||||| |||| |||||||||| |||||||||||||||||||||||||
Db      475 ACATTTCAAAGAT-GGCATTTCCCCCAATGAAATACACAAGTAAACATTC 523

```

RESULT 5

```

US-09-852-261-1
; Sequence 1, Application US/09852261
; Patent No. US20020083477A1
; GENERAL INFORMATION:
; APPLICANT: GOLDSPIK, GEOFFREY
; APPLICANT: TERENCE, GIORGIO
; TITLE OF INVENTION: REPAIR OF NERVE DAMAGE
; FILE REFERENCE: 117-351
; CURRENT APPLICATION NUMBER: US/09/852,261
; CURRENT FILING DATE: 2001-05-10
; PRIOR APPLICATION NUMBER: GB 0011278.9
; PRIOR FILING DATE: 2000-05-10
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 517
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-852-261-1

```

```

Query Match          60.3%; Score 325.2; DB 9; Length 517;
Best Local Similarity 81.2%; Pred. No. 1.4e-97;
Matches 429; Conservative 0; Mismatches 88; Indels 11; Gaps 4;

```

```

Qy      1 GGACCAGAGACCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTTCGTGTGTGGACCA 60
        ||||| ||||| || |||||||||||||||| |||||||||||||||| |||
Db      1 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTTCGTGTGTGGAGAC 60

Qy      61 AGGGGCTTTTACTTCAACAAGCCACAGTCTATGGCTCCAGCATTTCGGAGGGCACCACAG 120
        |||||||||| |||||||||||||||| |||||||||||||||| ||||||||
Db      61 AGGGGCTTTTATTTCAACAAGCCACAGGGTATGGCTCCAGCAGTTCGGAGGGCGCCTCAG 120

Qy      121 ACGGGCATTGTGGATGAGTGTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 180
        || ||||| |||||||||| |||||||||||||||| ||||||||||||||||
Db      121 ACAGGCATCGTGGATGAGTGTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 180

Qy      181 TGTGTCCGCTGCAAGCCTACAAAGTCAGCTCGTTCCATCCGGGCCCAGCGCCACACTGAC 240
        ||| ||| |||||| |||||||||| || |||| |||||||||| |||||
Db      181 TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCAGCGCCACACCGAC 240

```



```

Qy      181 TGTGTCCGCTGCAAGCCTACAAAGTCAGCTCGTTCCATCCGGGCCCAGCGCCCACTGAC 240
      || | | | | ||||| | ||||| || | ||||| ||||| |||||
Db      491 TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCAGCGCCACACCGAC 550

Qy      241 ATGCCCCAAGACTCAGAAGTCCCAGCCCCTATCGACACACAAGAAAAGGAAGCTGCAAAGG 300
      ||||| ||||| |||
Db      551 ATGCCCCAAGACCCAG----- 565

Qy      301 AGAAGGAAAGGAAGTACACTTGAAGAACACAAGTAGAGGAAGTGCAGGAAACAAGACCTA 360
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      566 -----AAGGAAGTACATTTGAAGAACGCAAGTAGAGGGAGTGCAGGAAACAAGAACTA 618

Qy      361 CAGAATGTAGGAGGAGCCTCCCGAGGAACAGAAAATGCCACGTCACCGCAAGATCCTTTG 420
      ||| ||||| || | ||||| ||||| | | || | ||||| |||||
Db      619 CAGGATGTAGGAAGACCCTCCTGAGGAGTGAAGAGTGACATGCCACCGCAGGATCCTTTG 678

Qy      421 --CTGCTTGAGCAACCTGCAAAACATCGGAACACCTGCCAAATATCAATAATGAGTTCAA 478
      |||| ||| ||||| |||| | ||||| |||| | ||||| |||||
Db      679 CTCTGCACGAGTTACCTGTTAAACTTTGGAACACCTACCA-----AAAAATAAGTTTGA 732

Qy      479 TATCATTTTCAAGATGGGCATTTCCCTCAATGAAATACACAAGTAAACATTCC 531
      || ||||| | ||||| ||||| ||||| ||||| ||||| |||||
Db      733 TAACATTTTAAAGATGGGCGTTTCCCCCAATGAAATACACAAGTAAACATTCC 785

```

# RESULT 7

US-09-880-107-3739

```

; Sequence 3739, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3739
; LENGTH: 7260
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 X57025
US-09-880-107-3739

```

```

Query Match          50.9%; Score 274.2; DB 10; Length 7260;
Best Local Similarity 76.0%; Pred. No. 5.7e-80;
Matches 405; Conservative 0; Mismatches 68; Indels 60; Gaps 3;

```



; LENGTH: 7260  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: Genbank Accession No. US20030134324A1 X57025  
US-09-873-319-707

Query Match 50.9%; Score 274.2; DB 13; Length 7260;  
Best Local Similarity 76.0%; Pred. No. 5.7e-80;  
Matches 405; Conservative 0; Mismatches 68; Indels 60; Gaps 3;

```
Qy      1 GGACCAGAGACCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTCGTGTGTGGACCA 60
      ||||| ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     311 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 370

Qy      61 AGGGGCTTTTACTTCAACAAGCCCACAGTCTATGGCTCCAGCATTCCGAGGGGCACCACAG 120
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     371 AGGGGCTTTTATTTCAACAAGCCCACAGGGTATGGCTCCAGCAGTCCGAGGGGCGCCTCAG 430

Qy     121 ACGGGCATTGTGGATGAGTGTGTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 180
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     431 ACAGGCATCGTGGATGAGTGTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 490

Qy     181 TGTGTCCGCTGCAAGCCTACAAAGTCAGCTCGTTCCATCCGGGCCCAGCGCCACACTGAC 240
      || | || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     491 TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGTCTGTCCGTGCCCAGCGCCACACCGAC 550

Qy     241 ATGCCCAAGACTCAGAAGTCCCAGCCCCTATCGACACACAAGAAAAGGAAGCTGCAAAGG 300
      ||||| ||||| |||
Db     551 ATGCCCAAGACCCAG----- 565

Qy     301 AGAAGGAAAGGAAGTACACTTGAAGAACACAAGTAGAGGAAGTGCAGGAAACAAGACCTA 360
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     566 -----AAGGAAGTACATTTGAAGAACGCAAGTAGAGGAGTGCAGGAAACAAGAACTA 618

Qy     361 CAGAATGTAGGAGGAGCCTCCCGAGGAACAGAAAATGCCACGTCACCGCAAGATCCTTTG 420
      ||| ||||| || ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     619 CAGGATGTAGGAAGACCCCTCCTGAGGAGTGAAGAGTGACATGCCACCGCAGGATCCTTTG 678

Qy     421 --CTGCTTGAGCAACCTGCAAAACATCGGAACACCTGCCAAATATCAATAATGAGTTCAA 478
      |||| ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     679 CTCTGCACGAGTTACCTGTTAAACTTTGGAACACCTACCA-----AAAAATAAGTTTGA 732

Qy     479 TATCATTTTCAGAGATGGGCATTTCCCTCAATGAAATACACAAGTAAACATTCC 531
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     733 TAACATTTAAAAGATGGGCGTTTCCCCCAATGAAATACACAAGTAAACATTCC 785
```

RESULT 9  
US-09-960-706-1066  
; Sequence 1066, Application US/09960706  
; Publication No. US20030134280A1  
; GENERAL INFORMATION:  
; APPLICANT: Munger, William E.  
; TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic Hyperplasia Using  
; TITLE OF INVENTION: Gene Expression Profiles

```

; FILE REFERENCE: 44921-5029-01US
; CURRENT APPLICATION NUMBER: US/09/960,706
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 60/223,323
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: 09/873,319
; PRIOR FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 1124
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1066
; LENGTH: 7260
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20030134280A1 X57025
US-09-960-706-1066

```

Qy		1 GGACCGAGAGACCCTTTGTCGGGGCCTTGAGCTGGTGGACGCTCTTCAGTTCTGTGTGGACCA	60
Db		311 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC	370
Qy		61 AGGGGCTTTTTACTTCAACAAGCCCAAGTCTATGGCTCCAGCATTCGGAGGGCACCAACAG	120
Db		371 AGGGGCTTTTTATTTCACAACAAGCCCAAGGGTATGGCTCCAGCAGTTCGGAGGGCGCCTCAG	430
Qy		121 ACGGGCATTGTGGATGAGTGTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC	180
Db		431 ACAGGCATCGTGGATGAGTGTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT	490
Qy		181 TGTGTCCGCTGCAAGCCTACAAAGTCAGCTCGTTCCATCCGGGCCAGCGCCACACTGAC	240
Db		491 TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCAGCGCCACACCGAC	550
Qy		241 ATGCCCCAAGACTCAGAAGTCCCAGCCCCCTATCGACACACAAGAAAAGGAAGCTGCAAAGG	300
Db		551 ATGCCCCAAGACCCAG-----	565
Qy		301 AGAAGGAAAGGAAGTACACTTGAAGAACACAAGTAGAGGAAGTGCAGGAAACAAGACCTA	360
Db		566 -----AAGGAAGTACATTTGAAGAACGCAAGTAGAGGGAGTGCAGGAAACAAGAACTA	618
Qy		361 CAGAATGTAGGAGGAGCCTCCCAGGGAACAGAAAATGCCACGTACCCGCAAGATCCTTTG	420
Db		619 CAGGATGTAGGAAGACCCCTCTGAGGAGTGAAGAGTGACATGCCACCGCAGGATCCTTTG	678
Qy		421 --CTGCTTGAGCAACCTGCAAAACATCGGAACACCTGCCAAATATCAATAATGAGTTCAA	478
Db		679 CTCTGCACGAGTTACCTGTTAACCTTTGGAACACCTACCA-----AAAAATAAGTTTGA	732
Qy		479 TATCATTTTCAGAGATGGGCATTTCCCTCAATGAAATACACAAGTAAACATTCC	531
Db		733 TAACATTTTAAAGATGGGCGTTTCCCCCCAATGAAATACACAAGTAAACATTCC	785



US-10-136-639-4

Query Match 50.9%; Score 274.2; DB 15; Length 7260;  
Best Local Similarity 76.0%; Pred. No. 5.7e-80;  
Matches 405; Conservative 0; Mismatches 68; Indels 60; Gaps 3;

[illegible]

[illegible]

RESULT 11

US-10-207-655-54

; Sequence 54, Application US/10207655

; Publication No. US20030118592A1

; GENERAL INFORMATION:

APPLICANT: Ledbetter, Jeffrey A.

APPLICANT: Hayden-Ledbetter, Martha S.

TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS

; FILE REFERENCE: 390069.401C1

CURRENT APPLICATION NUMBER: US/10/207,655

CURRENT FILING DATE: 2002-07-25

; NUMBER OF SEQ ID NOS: 426

```

; SOFTWARE: PatentIn version 3.0

```

; SEO ID NO 54

LENGTH: 725

; TYPE: DNA

ORGANISM: Homo sapiens

US-10-207-655-54

Query Match 50.6%; Score 272.6; DB 15; Length 725;

Best Local Similarity 75.8%; Pred. No. 5.6e-80;

Matches 404; Conservative 0; Mismatches 69; Indels 60; Gaps 3;

Qy	1	GGACCAGACCCCTTTGCGGGGTCTGAGCTGGTGGACGCTCTTCAGTTCGTGTGTGGACCA	60
Db	156	GGACCGGAGACGCTCTGCGGGGTCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC	215
Qy	61	AGGGGCTTTTTACTTCAACAAGCCCACAGTCTATGGCTCCAGCATTCGGAGGGCACCA CAG	120
Db	216	AGGGGCTTTTTATTTCACAAGCCCACAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAG	275
Qy	121	ACGGGCATTGTGGATGAGTGTTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC	180
Db	276	ACAGGTATCGTGGATGAGTGCTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT	335
Qy	181	TGTGTCCGCTGCAAGCCTACAAAGTCAGCTCGTTCCATCCGGGCCCCAGCGCCACACTGAC	240
Db	336	TGCGCACCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCAGCGCCACACCGAC	395
Qy	241	ATGCCCCAAGACTCAGAAGTCCCAGCCCCTATCGACACACAAGAAAAGGAAGCTGCAAAGG	300
Db	396	ATGCCCCAAGACCCAG-----	410
Qy	301	AGAAGGAAAGGAAGTACACTTGAAGAACACAAGTAGAGGAAGTGCAGGAAAACAAGACCTA	360
Db	411	-----AAGGAAGTACATTTGAAGAACGCAAGTAGAGGGAGTGCAGGAAAACAAGAACTA	463
Qy	361	CAGAATGTAGGAGGAGCCTCCCAGGAAACAGAAAAATGCCACGTCACCGCAAGATCCTTTG	420
Db	464	CAGGATGTAGGAAGACCCTCCTGAGGAGTGAAGAGTGACATGCCACCGCAGGATCCTTTG	523

```

Qy      421 --CTGCTTGAGCAACCTGCAAAACATCGGAACACCTGCCAAATATCAATAATGAGTTCAA 478
      |||| ||| |||| |||| | |||| |||| |||
Db      524 CTCTGCACGAGTTACCTGTTAAACTTTGGAACACCTACCA-----AAAAATAAGTTTGA 577

Qy      479 TATCATTTTCAGAGATGGGCATTTCCCTCAATGAAATACACAAGTAAACATTCC 531
      || |||| | |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Db      578 TAACATTTTAAAGATGGGCGTTTCCCCCAATGAAATACACAAGTAAACATTCC 630

```

```

RESULT 12
US-09-852-261-13
; Sequence 13, Application US/09852261
; Patent No. US20020083477A1
; GENERAL INFORMATION:
; APPLICANT: GOLDSPIK, GEOFFREY
; APPLICANT: TERENCE, GIORGIO
; TITLE OF INVENTION: REPAIR OF NERVE DAMAGE
; FILE REFERENCE: 117-351
; CURRENT APPLICATION NUMBER: US/09/852,261
; CURRENT FILING DATE: 2001-05-10
; PRIOR APPLICATION NUMBER: GB 0011278.9
; PRIOR FILING DATE: 2000-05-10
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 471
; TYPE: DNA
; ORGANISM: Oryctolagus cuniculus
US-09-852-261-13

```

```

Query Match          48.6%; Score 262; DB 9; Length 471;
Best Local Similarity 74.7%; Pred. No. 1.5e-76;
Matches 396; Conservative 0; Mismatches 75; Indels 59; Gaps 3;

```

```

Qy      1 GGACCAGAGACCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTCGTGTGTGGACCA 60
      |||| |||| || |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Db      1 GGACCGGAGACGCTCTGCGGTGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 60

Qy     61 AGGGGCTTTTACTTCAACAAGCCCACAGTCTATGGCTCCAGCATTCCGAGGGGCACCACAG 120
      |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Db     61 AGGGGCTTTTATTTCAACAAGCCCACAGGATACGGCTCCAGCAGTCCGAGGGGCACCTCAG 120

Qy    121 ACGGGCATTGTGGATGAGTGTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 180
      || |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Db    121 ACAGGCATCGTGGATGAGTGTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 180

Qy    181 TGTGTCCGCTGCAAGCCTACAAAGTCAGCTCGTTCCATCCGGGGCCAGCGCCACACTGAC 240
      |||| | | |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Db    181 TGTGCACCCCTCAAGCCGGCAAAGGCAGCCCGCTCCGTCCGTGCCAGCGCCACACCGAC 240

Qy    241 ATGCCCAAGACTCAGAAGTCCCAGCCCCCTATCGACACACAAGAAAAGGAAGCTGCAAAGG 300
      |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Db    241 ATGCCCAAGACTCAG----- 255

Qy    301 AGAAGGAAAGGAAGTACACTTGAAGAACAAGTAGAGGAAGTGACAGGAAACAAGACCTA 360
      |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||

```

```

Db      256 -----AAGGAAGTACATTTGAAGAACACAAGTAGAGGGAGTGCAGGAAACAAGAACTA 308
Qy      361 CAGAATGTAGGAGGAGCCTCCCGAGGAACAGAAAATGCCACGTCACCGCAAGATCCTTTG 420
      ||| ||||| || ||| | |||| | || | || | ||||| || |||||
Db      309 CAGGATGTAGGAAGACCCTTCTGAGGAGTGAAGAAGGACAGGCCACCGCAGGACCCTTTG 368
Qy      421 CTGCTTGAGCAACCTGCAAAACATCGGAACACCTGCCAAATATCAATAATGAGTTCAATA 480
      || | ||||| ||||| ||| ||| ||| || ||||| || ||
Db      369 CTCTGCACAGTTACCTGTAAACATTGGAATACCGGCCA-----AAAAATAAGTTTGATC 422
Qy      481 TCATTTTCAGAGATGGGCATTTCCCTCAATGAAATACACAAGTAAACATTTC 530
      ||||| |||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      423 ACATTTCAAAGAT-GGCATTTCCCCCAATGAAATACACAAGTAAACATTTC 471

```

# RESULT 13

US-10-251-661-7

; Sequence 7, Application US/10251661

; Publication No. US20030166555A1

; GENERAL INFORMATION:

; APPLICANT: Alberini, Cristina M.

; APPLICANT: Bear, Mark F.

; TITLE OF INVENTION: Methods and Compositions for Regulating

; TITLE OF INVENTION: Memory Consolidation

; FILE REFERENCE: 3499.1001-003

; CURRENT APPLICATION NUMBER: US/10/251,661

; CURRENT FILING DATE: 2002-09-20

; PRIOR APPLICATION NUMBER: 60/193,614

; PRIOR FILING DATE: 2000-03-31

; PRIOR APPLICATION NUMBER: PCT/US01/10661

; PRIOR FILING DATE: 2001-04-02

; NUMBER OF SEQ ID NOS: 12

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 7

; LENGTH: 612

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (151)...(564)

US-10-251-661-7

Query Match 41.6%; Score 224; DB 13; Length 612;

Best Local Similarity 75.6%; Pred. No. 8e-64;

Matches 316; Conservative 0; Mismatches 50; Indels 52; Gaps 1;

```

Qy      1 GGACCAGAGACCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTTCGTGTGTGGACCA 60
      ||||| ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      247 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTTCGTGTGTGGAGAC 306
Qy      61 AGGGGCTTTTACTTCAACAAGCCCACAGTCTATGGCTCCAGCATTTCGGAGGGCACCACAG 120
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      307 AGGGGCTTTTATTTCAACAAGCCCACAGGTATGGCTCCAGCAGTTCGGAGGGCGCCTCAG 366
Qy      121 ACGGGCATTGTGGATGAGTGTGCTTCCGGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 180
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      367 ACAGGCATCGTGGATGAGTGTGCTTCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 426

```



```

Db      198 CCTACAAAAGCAGCCCGCTCTATCCGTGCCAGCGCCACACTGACATGCCCAAGACTCAG 257
Qy      256 AAGTCCCAGCCCCTATC 272
          | | ||| | ||
Db      258 GCATGCAAGCTTGTCTC 274

```

RESULT 15

```

US-09-852-261-9
; Sequence 9, Application US/09852261
; Patent No. US20020083477A1
; GENERAL INFORMATION:
; APPLICANT: GOLDSPINK, GEOFFREY
; APPLICANT: TERENCE, GIORGIO
; TITLE OF INVENTION: REPAIR OF NERVE DAMAGE
; FILE REFERENCE: 117-351
; CURRENT APPLICATION NUMBER: US/09/852,261
; CURRENT FILING DATE: 2001-05-10
; PRIOR APPLICATION NUMBER: GB 0011278.9
; PRIOR FILING DATE: 2000-05-10
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 318
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-852-261-9

```

```

Query Match          38.4%; Score 206.8; DB 9; Length 318;
Best Local Similarity 87.6%; Pred. No. 3e-58;
Matches 226; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

```

```

Qy      1 GGACCAGAGACCCTTTGCGGGGCTGAGCTGGTGGACGCTCTTCAGTTCGTGTGTGGACCA 60
          ||||| ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1 GGACCGGAGACGCTCTGCGGGGCTGAGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGAC 60

Qy     61 AGGGGCTTTTACTTCAACAAGCCACAGTCTATGGCTCCAGCATTCCGAGGGGCACCACAG 120
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     61 AGGGGCTTTTATTTCAACAAGCCACAGGGTATGGCTCCAGCAGTCCGAGGGGCGCCTCAG 120

Qy    121 ACGGGCATTGTGGATGAGTGTGCTTCCGAGCTGTGATCTGAGGAGGCTGGAGATGTAC 180
          || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    121 ACAGGCATCGTGGATGAGTGTGCTTCCGAGCTGTGATCTAAGGAGGCTGGAGATGTAT 180

Qy    181 TGTGTCCGCTGCAAGCCTACAAAGTCAGCTCGTTCATCCGGGCCCAGCGCCACACTGAC 240
          || | | | ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    181 TGCGCACCCCTCAAGCCTGCCAAGTCAGCTCGCTCTGTCCGTGCCAGCGCCACACCGAC 240

Qy    241 ATGCCCAAGACTCAGAAG 258
          ||||| |||||
Db    241 ATGCCCAAGACCCAGAAG 258

```

```

Search completed: December 13, 2003, 11:56:46
Job time : 241.655 secs

```